Harmonic analysis approach to Voiculescu's R-transform

A Thesis Submitted to the
College of Graduate Studies and Research
in Partial Fulfillment of the Requirements
for the degree of Master of Science
in the Department of Mathematics and Statistics
University of Saskatchewan
Saskatoon

By
Pei-Lun Tseng

Permission to Use

In presenting this thesis in partial fulfilment of the requirements for a Postgraduate degree from the University of Saskatchewan, I agree that the Libraries of this University may make it freely available for inspection. I further agree that permission for copying of this thesis in any manner, in whole or in part, for scholarly purposes may be granted by the professor or professors who supervised my thesis work or, in their absence, by the Head of the Department or the Dean of the College in which my thesis work was done. It is understood that any copying or publication or use of this thesis or parts thereof for financial gain shall not be allowed without my written permission. It is also understood that due recognition shall be given to me and to the University of Saskatchewan in any scholarly use which may be made of any material in my thesis.

Requests for permission to copy or to make other use of material in this thesis in whole or part should be addressed to:

Head of the Department of Mathematics and Statistics
Room 142 McLean Hall
106 Wiggins Road
University of Saskatchewan
Saskatoon, Saskatchewan
Canada
S7N 5E6

Abstract

In this thesis, we present a study of Voiculescu's R-transform from the complex and harmonic analysis point of view. This study is based on the unpublished notes [4] of James Mingo and Roland Speicher, in which the main feature is the phenomenon of analytic subordination. Thus, this thesis is not an original research, but author's own understanding (of the aforementioned book [4]) and comments on this particular topic in free harmonic analysis.

Acknowledgements

To start with, I would like to express my appreciation to the Department of Mathematics and Statistics at the University of Saskatchewan, in particular Dr. Raj Srinivasan, for providing me the opportunity to do my thesis Harmonic Analysis Approach to Voiculescu's R-transform.

I would like to extend my special gratitude to my supervisor, Dr. Jiun-Chau Wang. He invited me to visit the University of Saskatchewan in 2013 and further offered me a MSc position to do my research of interest. He has kindly supported me academically and financially for my research in free probability. Dr. Wang assisted me to clarify mathematical concepts in our regular meetings, which gained me the great understandings in the field. Not only my schoolwork, he also cared about my personal life. I could not have accomplished the thesis without his supports.

I also appreciate the support from Dr. Jacek Szmigielski, who was inspirational and gave me encouragements to overcome challenges.

Finally, throughout the course of two-year study, I have consistently received generous supports from many of my colleagues in the department. I would like to thank Jessica, Yun-Yang Wang, Dong Yue and all other faculty members, staffs, and classmates in the Department of Mathematics and Statistics for their assistances in various regards.

Contents

Chapter 1

Introduction

A *random matrix* is simply a matrix of random variables. Thus, let $X_n = [a_{ij}]$ be a $n \times n$ random matrix, where each a_{ij} is a complex-valued random variable defined on some probability space $(\Omega, \mathfrak{F}, P)$. The *empirical spectral distribution* μ_{X_n} of X_n is the random probability measure defined by

$$\mu_{X_n} = \frac{1}{n}\sum_{j=1}^{n} \delta_{\lambda_j},$$

where $\lambda_j = \lambda_j(\omega)$ $(1 \le j \le n)$ are the eigenvalues of the numerical matrix $X_n(\omega) = [a_{ij}(\omega)]$ for $\omega \in \Omega$. (Here the multiplicity of each eigenvalue is counted in the realization $X_n(\omega)$.) The knowledge of the (random) spectrum of X_n is equivalent to that of the measure μ_{X_n}.

The random matrix ensemble X_n is said to be *selfadjoint* (resp., *symmetric*) if each realization $X_n(\omega)$ is a selfadjoint (resp., symmetric) matrix. In both cases, the empirical spectral distribution μ_{X_n} is supported on the real line $\mathbb{R}$. Moreover, given a probability measure ν on $\mathbb{R}$, we say that X_n *converges weakly in probability* to ν (and denote this by $X_n \Rightarrow \nu$) if for any given $\varepsilon > 0$ and for any bounded and continuous function $f : \mathbb{R} \to \mathbb{R}$, one has

$$\lim_{n\to\infty} P(\ \omega \in \Omega \mid |\int_{\mathbb{R}} f d\mu_{X_n} - \int_{\mathbb{R}} f d\nu| > \varepsilon) = 0.$$

To simplity our discussion, we focus on the symmetric case in which the variable a_{ij} takes values in $\mathbb{R}$. In addition to the symmetric structure on X_n, we introduce further some independence and moment conditions to the entries a_{ij}. We assume that (i) the family $\{a_{ii} : i \in \mathbb{N}\}$ of diagonal entries is independent and identically distributed (i.i.d.) and so is the family $\{a_{ij} : i, j \in \mathbb{N}, i < j\}$ of variables in the upper diagonal part of our matrix ensemble, (ii) the diagonal family $\{a_{ii} : i \in \mathbb{N}\}$ is independent from the upper diagonal family $\{a_{ij} : i, j \in \mathbb{N}, i < j\}$, (iii) $Ea_{ij} = 0$ and $Ea_{ij}^2 = 1$. We call such a random matrix X_n a *Wigner matrix*. Under the above assumptions, Wigner proved that

$$\frac{1}{\sqrt{n}} X_n \Rightarrow \gamma,$$

where the limit distribution γ is the semicircle law defined by

$$d\gamma(t) = \frac{1}{2\pi}\sqrt{4 - t^2}dt, \quad t \in [-2, 2].$$

(See [5] for a proof of Wigner's result.)

The preceding result shows that the spectrum of $n^{-1/2}X_n$ is determined asymptotically by a deterministic (i.e., non-random) distribution in high probability as the dimension $n \to \infty$. After knowing this result, we ask a natural question as follows.

Problem 1. *What can we say about the asymptotic behavior of the spectrum for any reasonable function* $f\left(n^{-1/2}X_n^{(1)}, n^{-1/2}X_n^{(2)}, \cdots, n^{-1/2}X_n^{(p)}\right)$ *in several such Wigner random matrices?*

For examples, the function f could be as simple as the elementary polynomials like $f(x_1, x_2) = x_1 + x_2$ or $f(x_1, x_2) = x_1x_2$ for $p = 2$.

While the general case of the problem remains open to date, Voiculescu gave an answer to the case $f(x_1, x_2) = x_1 + x_2$ in his 1991 paper [6]. More precisely, consider two Hermitian random matrix ensembles X_n and Y_n satisfying the following conditions:

1. All entries of X_n are independent to that of Y_n,

2. For any $n\times n$ unitary matrix U, the joint probability distribution of the entries in the unitary conjugation UY_nU^* is the same as the joint probability distribution of the original entries in Y_n,

3. There exist two deterministic probability distributions μ and ν such that $X_n \Rightarrow \mu$ and $Y_n \Rightarrow \nu$.

Then Voiculescu showed that the sum $f(X_n, Y_n) = X_n + Y_n$ converges weakly in probability to a deterministic probability distribution $\mu \boxplus \nu$, the *free convolution* of the measures μ and ν (See Chapter 3 for the construction of the free convolution.) In other words, we can now understand the asymptotics of the spectrum of $X_n + Y_n$ by studying the measure $\mu \boxplus \nu$. In particular, if we assume that both X_n and Y_n are Wigner matrices of Gaussian entries (in order to satisfy the unitary invariance condition (2)), then we know in high probability that the spectrum of any realization of $X_n + Y_n$ is well approximated by the free convolution $\gamma \boxplus \gamma$ of the semicircle law with itself. (This free convolution turns out to be another semicircle law, but with variance 2.) Since Voiculescu discovered this connection between free probability and random matrices, free probability theory has become an interesting subject to various communities of scientists who use large random matrices in their models.

In this thesis we focus on the study of the free convolution $\mu \boxplus \nu$. Indeed, the measure $\mu \boxplus \nu$ is defined as the spectral measure of $a + b$, where a and b are two self-adjoint (non-commutative) random variables that are *free* and having distributions μ and ν, respectively. (See Chapter 3 for details.) The computation of $\mu \boxplus \nu$ relies on the addition formula of Voiculescu's *R-transform*, namely,

$$R_{a+b} = R_a + R_b. \tag{1}$$

This is in parallel to the multiplicative property of the Fourier transform in classical probability, which serves as a main tool for analyzing the classical convolution of probability measures. We will present the construction of R-transform and a proof of (1). Moreover, we study analytic properties of the R-transform and its connections with the underlying distribution. This approach to R-transform uses complex and harmonic analysis methods. It is largely based on the unpublished book [4] of James Mingo and Roland Speicher and on a series of papers [8,9,10] in free harmonic analysis.

We emphasize that this thesis is not an original research, but our own summary and comments of the free harmonic analysis chapters in the book [4].

The organization of this thesis is as follows. In Chapter 2, we review the basics of non-commutative probability theory and the free product construction for C^*-probability spaces. Some useful properties of Cauchy transform of measures will also be reviewed here. Chapter 3 contains the proof of (1) and that of the existence of free convolution for probability distributions with bounded support. Free convolution for distributions with finite variance is discussed in Chapter 4. Finally, we show that the formula (1) holds for general probability distributions in Chapter 5.

Chapter 2

Preliminaries

2.1 Basics and Notations in Free Probability

In a purely algebraic framework, a *non-commutative probability space* $(\mathcal{A}, \varphi)$ consists of a unital algebra $\mathcal{A}$ over $\mathbb{C}$ and a linear functional $\varphi : \mathcal{A} \to \mathbb{C}$ with $\varphi(1) = 1$. Thus, the algebra $\mathcal{A}$ should be thought of as the algebra of (not necessarily commuting) random variables that contains the multiplicative identity 1 such that $1a = a1 = a$ for all $a \in \mathcal{A}$ and the functional φ as the expectation on these variables. The fact that these definitions carry a probabilistic flavour can be seen from the identity $\varphi(1) = 1$.

The algebra $\mathcal{A}$ is said to be a $*$*-algebra* if there exists an operation $* : \mathcal{A} \to \mathcal{A}$ such that (1) $(x+y)^* = x^* + y^*$, (2) $(xy)^* = y^* x^*$, (3) $1^* = 1$, (4) $(x^*)^* = x$, and (5) $(\lambda x)^* = \bar{\lambda} x^*$ for all $x, y \in \mathcal{A}$ and $\lambda \in \mathbb{C}$. If $\mathcal{A}$ is a unital $*$-algebra, and we have $\varphi(a^* a) \geq 0$ for all $a \in \mathcal{A}$, then we say that the functional φ is *positive*, and that $(\mathcal{A}, \varphi)$ is a $*$*-probability space*. If an element $a \in \mathcal{A}$ has the property $a^* = a$, then we say that a is *self-adjoint*. The expectation functional φ is said to be *faithful* if $\varphi(a^* a) = 0$ implies $a = 0$. If $\varphi(ab) = \varphi(ba)$ for all $a, b \in \mathcal{A}$, then we call φ a *trace*.

In the following definition, $\mathcal{A}_i$ being a *unital subalgebra* of $\mathcal{A}$ means that $\mathcal{A}_i$ is a subalgebra of $\mathcal{A}$ and $\mathcal{A}_i$ contains the multiplicative identity 1 of $\mathcal{A}$.

Definition 2. Let $(\mathcal{A}, \varphi)$ be a non-commutative probability space and I be an index set.

(1) For each $i \in I$, let $\mathcal{A}_i \subset \mathcal{A}$ be a unital subalgebra. The family $\{\mathcal{A}_i : i \in I\}$ is said to be *freely independent* (or, just *free* for short) , if we have

$$\varphi(a_1 a_2 \dots a_k) = 0$$

whenever $k \in \mathbb{N}$, $i(j) \in I$ $(1 \leq j \leq k)$, $i(1) \neq i(2) \neq \dots \neq i(k-1) \neq i(k)$, $a_j \in \mathcal{A}_{i(j)}$, and $\varphi(a_j) = 0$ for all $1 \leq j \leq k$.

(2) A set $\{a_i \in \mathcal{A} : i \in I\}$ of random variables is free if the unital algebras $\mathcal{A}_i = alg(1, a_i)$ generated by a_i $(i \in I)$ form a free family.

(3) In the context of $*$-probability space, random variables a_i, $i \in I$, are said to be free if the unital $*$-algebras $\mathcal{A}_i = alg(1, a_i, a_i^*)$ generated by a_i are free.

For $a \in (\mathcal{A}, \varphi)$, we write $a^\circ = a - \varphi(a)1$.

Example 3. *Let $a, b \in (\mathcal{A}, \varphi)$ be two free random variables.*

(1) By the definition of freeness, we have $\varphi(a^\circ b^\circ) = 0$. On the other hand, note that

$$\varphi(a^\circ b^\circ) = \varphi((a - \varphi(a)1)(b - \varphi(b)1)) = \varphi(ab) - \varphi(a)\varphi(b).$$

So, we have the result

$$\varphi(ab) = \varphi(a)\varphi(b). \tag{2.1}$$

(2) Next, we have $\varphi(a^\circ b^\circ a^\circ) = 0$ by definition and $\varphi(a^\circ b^\circ a^\circ) = \varphi(aba) - \varphi(a^2)\varphi(b) + 2\varphi(a)^2\varphi(b) - 2\varphi(a)\varphi(ba) = \varphi(aba) - \varphi(a^2)\varphi(b)$ by straightforward calculations and the equation (2.1). Therefore, we conclude that

$$\varphi(aba) = \varphi(a^2)\varphi(b). \tag{2.2}$$

(3) Similarly, we have $\varphi(a^\circ b^\circ a^\circ b^\circ) = 0$ by definition and

$$\varphi(a^\circ b^\circ a^\circ b^\circ) = \varphi(abab) - \varphi(a^2)\varphi(b)^2 - \varphi(a)^2\varphi(b^2) + \varphi(a)^2\varphi(b)^2,$$

implying that $\varphi(abab) = \varphi(a^2)\varphi(b)^2 + \varphi(a)^2\varphi(b^2) - \varphi(a)^2\varphi(b)^2$. In addition, if a commutes b then it would follow that

$$\varphi(abab) = \varphi(a^2b^2) = \varphi(a^2)\varphi(b^2).$$

Hence, we have the product of their variances

$$\varphi(a^\circ)\varphi(b^\circ) = 0.$$

In particular, this shows that two classical random variables cannot exhibit freeness, unless one of them is a constant (and hence deterministic) random variable.

Lemma 4. *Let $(\mathcal{A}, \varphi)$ be a non-commutative probability space, and let $(\mathcal{A}_i)_{i\in I}$ be a freely independent family of unital subalgebras of $\mathcal{A}$. Let $k, l \in \mathbb{N}$ and $i(1), \ldots, i(k), j(1), \ldots, j(l)$ be indices in I with*

$$i(1) \neq i(2) \neq \ldots \neq i(k) \text{ and } j(1) \neq j(2) \neq \ldots \neq j(l)$$

and let $a_s \in \mathcal{A}_{i(s)},\ b_t \in \mathcal{A}_{j(t)}$ such that $\varphi(a_s) = \varphi(b_t) = 0$ for $1 \le s \le k,\ 1 \le t \le l$. Then, we have

$$\varphi(a_1 \ldots a_k b_l \ldots b_1) = \begin{cases} \varphi(a_1 b_1) \ldots \varphi(a_k b_k) & \text{if } k = l, \text{ and } i(1) = j(1), \ldots, i(k) = j(k) \\ 0 & \text{otherwise.} \end{cases}$$

Proof. One has to iterate the following observation: either we have $i(k) \neq j(l)$, in whcih case

$$\varphi(a_1 \ldots a_k b_l \ldots b_1) = 0$$

by freeness, or we have $i(k) = j(1)$, which implies that

$$\begin{aligned} \varphi(a_1 \ldots a_k b_l \ldots b_1) &= \varphi(a_1 \ldots a_{k-1}[(a_k b_l)^\circ + \varphi(a_k b_l)1] b_{l-1} \ldots b_1) \\ &= \varphi((a_k b_l)^\circ)\varphi(a_1 \ldots a_{k-1} b_{l-1} \ldots b_1) + \varphi(a_k b_l)\varphi(a_1 \ldots a_{k-1} b_{l-1} \ldots b_1) \\ &= 0 + \varphi(a_k b_l)\varphi(a_1 \ldots a_{k-1} b_{l-1} \ldots b_1). \end{aligned}$$

□

Proposition 5. *Let $(\mathcal{A}, \varphi)$ be a non-commutative probability space, let $(\mathcal{A}_i)_{i\in I}$ be a freely independent family of unital subalgebras of $\mathcal{A}$, and let B be the subalgebra of $\mathcal{A}$ generated by $\bigcup_{i\in I}\mathcal{A}_i$. If $\varphi|_{\mathcal{A}_i}$ is a trace for $i \in I$, then $\varphi|_B$ is a trace.*

Proof. We observe that every element of B can be written as a linear combination of 1 and elements of the form $a_1 \ldots a_n$ for $n \in \mathbb{N}, i(1), \ldots, i(n) \in I$ and $i(1) \neq \ldots \neq i(n)$, $a_p \in \mathcal{A}_{i(p)}$ for all $1 \leq p \leq n$. Since $a_i = a_i^\circ + \varphi(a_i)1$, we may assume that $\varphi(a_i) = 0$ for $i = 1, 2, \ldots, n$. It suffices to prove the assertion for a and b of the special form $a = a_1 \ldots a_k$ and $b = b_l b_{l-1} \ldots b_1$ with $a_p \in \mathcal{A}_{i(p)}$ and $b_q \in \mathcal{A}_{j(q)}$, where $1 \leq p \leq k, 1 \leq q \leq l, i(1), \ldots, i(k), j(1), \ldots, j(l) \in I, i(1) \neq \ldots \neq i(k)$, $j(1) \neq \ldots \neq j(l)$, and such that $\varphi(a_1) = \ldots = \varphi(a_k) = \varphi(b_1) = \ldots \varphi(b_l) = 0$. But in this situation we can apply Lemma 4 and get

$$\begin{aligned}
\varphi(ab) = \varphi(a_1 \ldots a_k b_l \ldots b_1) &= \begin{cases} \varphi(a_1b_1)\ldots\varphi(a_kb_k) & \text{if } k = l, i(1) = j(1), \ldots i(k) = j(k) \\ 0 & \text{otherwise,} \end{cases} \\
\varphi(ba) = \varphi(b_l \ldots b_1 a_1 \ldots a_k) &= \begin{cases} \varphi(b_la_l)\ldots\varphi(b_1a_1) & \text{if } k = l, i(1) = j(1), \ldots, i(k) = j(k) \\ 0 & \text{otherwise.} \end{cases}
\end{aligned}$$

Since φ is trace on $\mathcal{A}_i$ for each $i \in I$, we conclude that $\varphi(ab) = \varphi(ba)$. □

Let $\mathcal{A}_1, \mathcal{A}_2$ be two subalgebras in $(\mathcal{A}, \varphi)$. If $\mathcal{A}_1, \mathcal{A}_2$ are free, we have the following result:

Theorem 6. *If $a_1, a_2, \ldots, a_k \in \mathcal{A}_1$, $b_1, b_2, \ldots,$ $b_k \in \mathcal{A}_2$, then $\varphi(a_1b_1a_2b_2\ldots a_kb_k)$ can be written as a sum of terms of the form*

$$\varphi(a_{i_1}), \ldots, \varphi(a_{i_r}), \varphi(b_{j_1}), \ldots, \varphi(b_{j_t}), \varphi(c_1d_1c_2d_2, \ldots c_ld_l)$$

where $i_1, \ldots, i_r, j_1 \ldots, j_t \in \{1, 2, \ldots, k\}$, $l < k$, and $c_1, \ldots c_l \in \mathcal{A}_1, d_1 \ldots, d_l \in \mathcal{A}_2$.

Proof. We prove this theorem by the mathematical induction. The case $k = 1$, which is the result of Lemma 4. Suppose that for the case $k < n$, the statement is true. Now, consider the case $k = n$. Let $S = a_1b_1 \ldots a_{n-1}b_{n-1}$, then

$$\begin{aligned}
\varphi(a_1b_1 \ldots a_nb_n) &= \varphi(Sa_nb_n) = \varphi(S(a_n^\circ + \varphi(a_n))(b_n^\circ + \varphi(b_n))) \\
&= \varphi(Sa_n^\circ b_n^\circ) + \varphi(a_n)\varphi(Sb_n^\circ) + \varphi(b_n)\varphi(Sa_n^\circ) + \varphi(a_n)\varphi(b_n)\varphi(S).
\end{aligned}$$

Note that $S = (a_1^\circ + \varphi(a_1))(b_1^\circ + \varphi(b_1)) \ldots (a_{n-1}^\circ + \varphi(a_{n-1}))(b_{n-1}^\circ + \varphi(b_{n-1}))$. Then, by freeness,

$$\varphi(a_1^\circ b_1^\circ \ldots a_n^\circ b_n^\circ) = 0 \text{ and } \varphi(a_1^\circ b_1^\circ \ldots b_{n-1}^\circ a_n^\circ) = 0.$$

In addition, $b_{n-1}b_n^\circ \in \mathcal{A}_2$. Hence, the case $k = n$ is also true. By the mathematical induction, we complete the proof. □

Now, we observe that

$$(a + b)^n = \sum_{i_1+\ldots+i_k+j_1+\ldots+j_k=n} \varphi(a^{i_1}b^{j_1} \ldots a^{i_k}b^{j_k})$$

where $i_1, i_2, \ldots, i_k, j_1, j_2, \ldots, j_k$ are non-negative integers. Thus, we have the following Corollary:

Corollary 7. *If a, b are free in a non-commutative probability space, then for $n \in \mathbb{N}$, $\varphi((a+b)^n) = P_{2n}(\varphi(a), \varphi(a^2) \ldots, \varphi(a^n), \varphi(b), \varphi(b^2) \ldots, \varphi(b^n))$ where P_{2n} is a polynomial in $2n$ variables.*

Proof. Apply Theorem 6 to the case of $\mathcal{A}_1 = alg(1, a)$ and $\mathcal{A}_2 = alg(1, b)$. $\square$

2.2 Construction of Free Product of ∗-Probability Spaces

Given $(\mathcal{A}_i, \varphi_i)_{i \in I}$ be a family of non-commutative probability spaces. For each $i \in I$, we let $\mathcal{A}_i^\circ = ker(\varphi_i)$. For every $n \in \mathbb{N}$ and $i(1), \ldots, i(n) \in I$ such that $i(1) \neq i(2) \neq \ldots \neq i(n)$, we set

$$W_{i(1),\ldots,i(n)} = span\{a_1 \ldots a_n | a_1 \in \mathcal{A}_{i(1)}^\circ, \ldots a_n \in \mathcal{A}_{i(n)}^\circ\},$$

where the span of the non-rediciuible words of letters over $\mathbb{C}$. We introduce the vector addition and scalar multiplication on $W_{i(1),\ldots,i(n)}$ so that $W_{i(1),\ldots i(n)}$ becomes a vector space over $\mathbb{C}$. Given $a, b \in W_{i(1),\ldots,i(n)}$ and $\gamma \in \mathbb{C}$, $a + b$ and $\gamma \cdot a$ is defined as follows:

$$\begin{aligned} a + b &\equiv \sum_{finite} \alpha_{a_1 \ldots a_n} a_1 \ldots a_n + \sum_{finite} \beta_{b_1 \ldots b_n} b_1 \ldots b_n, \\ \gamma \cdot a &\equiv \sum_{finite} (\gamma \alpha_{a_1 \ldots a_n}) a_1 \ldots a_n, \end{aligned}$$

where

$$a = \sum_{finite} \alpha_{a_1 \ldots a_n} a_1 \ldots a_n, \text{and } b = \sum_{finite} \beta_{b_1 \ldots b_n} b_1 \ldots b_n \quad (a_j, b_j \in \mathcal{A}_{i(j)}^\circ \text{ for } 1 \leq j \leq n,\ \alpha_{a_1 \ldots a_n}, \beta_{b_1 \ldots b_n} \in \mathbb{C}.)$$

Then, we consider the following set:

$$\mathcal{A} = \mathbb{C}1 \oplus \bigoplus_{n=1}^{\infty} \bigoplus_{\substack{i(1),\ldots,i(n) \in I; \\ i(1) \neq \ldots \neq i(n)}} W_{i(1),\ldots,i(n)}, \tag{2.3}$$

where 1 is just a symbol. Observe that $\mathbb{C}1 = \{r1 : r \in \mathbb{C}\}$ is a vector space obviously, and hence $\mathcal{A}$ is the direct sum of vector spaces, which implies that $\mathcal{A}$ is also a vector space.

We introduce the multiplication on $\mathcal{A}$ so that it becomes a unital algebra. Given $\alpha, \beta \in \mathbb{C}$,

$$a = a_1 a_2 \ldots a_s \text{ and } a' = a_1' a_2' \ldots a_t'$$

where $a_1 \in \mathcal{A}_{i(1)}^\circ, \ldots, a_s \in \mathcal{A}_{i(s)}^\circ$ and $a_1' \in \mathcal{A}_{j(1)}^\circ, \ldots, a_t' \in \mathcal{A}_{j(t)}^\circ$ with $i(1), \ldots,\ i(s), j(1), \ldots, j(t) \in I$ and $i(1) \neq i(2) \neq \ldots \neq i(s),\ j(1) \neq j(2) \neq \ldots \neq j(t)$. $\alpha 1 \cdot \beta 1$ is defined by $(\alpha\beta)1$, and so is $\beta 1 \cdot \alpha 1$. $\alpha 1 \cdot a$ and $a \cdot \alpha 1$ are both defined by $(\alpha) a_1 a_2 \ldots a_s$. The multiplication of $a \cdot a'$ can proceed by induction on $s + t$. If $i(s) \neq j(1)$, then $a \cdot a'$ is simply defined as $a_1 a_2 \ldots a_s a_1' a_2' \ldots a_t'$. If $i(s) = j(1)$, we note that $a_s a_1' = \varphi_{i(s)}(a_s a_1') + (a_s a_1')^\circ$, and consider the element $(a_1 \ldots a_{s-1}) \cdot (a_2' \ldots a_t')$ which is defined by the induction hypothesis. Then define the product of a and a' to be

$$a \cdot a' = a_1 a_2 \ldots a_{s-1} (a_s a_1')^\circ a_2' \ldots a_t' + \varphi_{i(s)}(a_s a_1')(a_1 \ldots a_{s-1}) \cdot (a_2' \ldots a_t').$$

Now, for elements $a, a' \in \mathcal{A}$, say $a = \alpha 1 + \sum_{finite} \alpha_i a_{i(1)} \dots a_{i(k)},\ a' = \beta 1 + \sum_{finite} \beta_j a'_{j(1)} \dots a'_{j(s)}$. The product of a and a' is defined as

$$a \cdot a' = (\alpha\beta)1 + \sum_{finite} \alpha\beta_j a'_{j(1)} \dots a'_{j(s)} + \sum_{finite} \beta\alpha_i a_{i(1)} \dots a_{i(k)} + \sum_{finite} \alpha_i\beta_j a_{i(1)} \dots a_{i(k)} a'_{j(1)} \dots a'_{j(s)}.$$

Then $a \cdot a' \in \mathcal{A}$, and it is easy to verify that it satisfies the associative law and distributive law. Hence, $\mathcal{A}$ is an algebra. Note that for each $i \in I$, the algebra $\mathcal{A}_i$ is naturally embedded inside $\mathcal{A}$ via $\mathcal{A}_i = \mathbb{C}1 \oplus W_i$ (we identify the unit $1 = 1_{\mathcal{A}_i}$.) Hence, for each $i \in I$, $\mathcal{A}_i$ can be regarded as a subalgebra of $\mathcal{A}$.

$\mathcal{A}$ is called the free product of algebras $(\mathcal{A}_i)_{i \in I}$, and $\mathcal{A}$ is denoted by $*_{i \in I}\mathcal{A}_i$. The free product of the linear functionals $(\varphi)_{i \in I}$ is defined as the unique linear functional $\varphi : \mathcal{A} \to \mathbb{C}$ which satisfies $\varphi(1) = 1$ and $\varphi(a) = 0$ if $a \in W_{i(1),\dots,i(s)}$ for every $s \in \mathbb{N}$ and every $i(1), \dots i(s) \in I$ are such that $i(1) \neq i(2) \neq \dots \neq i(s)$. The notation used for this functional φ is $*_{i \in I}\varphi_i$. Then the corresponding non-commutative probability space $(\mathcal{A}, \varphi)$ is called the *free product of non-commutative probability spaces* $(\mathcal{A}_i, \varphi_i)$, for $i \in I$, and one writes

$$(\mathcal{A}, \varphi) = *_{i \in I}(\mathcal{A}_i, \varphi_i).$$

Thus, it is obvious that $(\mathcal{A}_i)_{i \in I}$ are free in $(\mathcal{A}, \varphi)$. On the other hand, we observe that for $b \in \mathcal{A}_i$

$$\varphi(b) = \varphi(b^\circ + \varphi_i(b)) = \varphi(b^\circ) + \varphi_i(b)1 = \varphi_i(b).$$

Hence, we have the following result for free product of non-commutative probability spaces:

Theorem 8. *Let $(\mathcal{A}_i, \varphi_i)_{i \in I}$ be a family of non-commutative probability spaces, and let $(\mathcal{A}, \varphi)$ be their free product. Then $\mathcal{A}_i \subset \mathcal{A},\ \varphi|_{\mathcal{A}_i} = \varphi_i$ for all $i \in I$ and $(\mathcal{A}_i)_{i \in I}$ are free in $(\mathcal{A}, \varphi)$.*

Proposition 9. *Let $(\mathcal{A}_i, \varphi_i)_{i \in I}$ be a family of non-commutative probability spaces, and let $(\mathcal{A}, \varphi)$ be their free product. If φ_i is a trace for $i \in I$ then φ is a trace on $\mathcal{A}$.*

Proof. Since $(\mathcal{A}_i)_{i \in I}$ are free and $\varphi|_{\mathcal{A}_i} = \varphi_i$ is a trace for each $i \in I$, Proposition 5 implies that φ is a trace on the subalgebra generated by $\cup_{i \in I}\mathcal{A}_i$ which is all of $\mathcal{A}$. □

If $(\mathcal{A}_i, \varphi_i)_{i \in I}$ are not only non-commutative probability spaces, but $*$-probability spaces. We define $\mathcal{A}$ as (2.3), and then we have the $*$-operation on $\mathcal{A}$ maps $W_{i(1),\dots,i(n)}$ onto $W_{i(1),\dots,i(n)}$, via the formula

$$(a_1 \dots a_k)^* = a_k^* \dots a_1^*$$

where $a_j \in \mathcal{A}^\circ_{i(j)},\ 1 \leq j \leq k,\ i(1), \dots, i(k) \in I$ are such that $i(1) \neq \dots \neq i(k)$. This immediately implies that $\varphi(a^*) = \overline{\varphi(a)}$ for all $a \in \mathcal{A}$ where $\varphi = *_{i \in I}\varphi_i$.

Theorem 10. *Let $(\mathcal{A}_i, \varphi_i)_{i \in I}$ be a family of $*$-probability spaces. Then the corrpesonding free product $(\mathcal{A}, \varphi)$ is also a $*$-probability space; that is, φ is positive.*

In order to prove Theorem 10, we review some results of positive matrices. Note that for a matrix $A \in M_n(\mathbb{C})$, the following are equivalent:

(1) A is positive;

(2) A is self-adjoint and all its eigenvalues are in $[0, \infty)$;

(3) A can be written in the form $A = X^*X$ for some $X \in M_n(\mathbb{C})$;

(4) $\langle A\xi, \xi \rangle \geq 0$ for $\xi \in M_n(\mathbb{C})$.

Definition 11. Given $A = (a_{ij})$ and $B = (b_{ij})$ in $M_n(\mathbb{C})$, the *Schur product* of A and B is the matrix $S = (a_{ij}b_{ij})_{i,j=1}^n$.

Lemma 12. *Let $A = (a_{ij})$ and $B = (b_{ij})$ be two positive matrices in $M_n(\mathbb{C})$. Then, the Schur product $S = (a_{ij}b_{ij})$ is a positive matrix.*

Proof. For $\xi = (\xi_1, \xi_2, \ldots, \xi_n) \in \mathbb{C}^n$, we have

$$\langle S\xi, \xi \rangle = \sum_{i,j=1}^{n} a_{ij} b_{ij} \xi_j \overline{\xi_i}.$$

Since A is positive, $A = X^*X$ where $X = (x_{ij}) \in M_n(\mathbb{C})$. Hence, for $1 \leq i, j \leq n$, we have $a_{ij} = \sum\limits_{k=1}^{n} \overline{x_{ki}} x_{kj}$. Thus,

$$\langle S\xi, \xi \rangle = \sum_{i,j,k=1}^{n} \overline{x_{ki}} x_{kj} b_{ij} \xi_j \overline{\xi_i} = \sum_{k=1}^{n} \sum_{i,j=1}^{n} b_{ij} (\xi_j x_{kj}) (\overline{\xi_i x_{ki}}) = \sum_{k=1}^{n} \langle B\eta_k, \eta_k \rangle \geq 0$$

where $\eta_k = (\xi_1 x_{k1}, \ldots, \xi_n x_{kn}) \in \mathbb{C}^n$ for $1 \leq k \leq n$. $\square$

Lemma 13. *If $\mathcal{A}$ is a unital $*$-algebra equipped with a linear functional $\varphi : \mathcal{A} \to \mathbb{C}$, then the following statements are equivalent:*

(1) φ is positive;

(2) for all $n \in \mathbb{N}$, and $a_1, \ldots, a_n \in \mathcal{A}$, the matrix $(\varphi(a_i^* a_j))_{i,j=1}^n \in M_n(\mathbb{C})$ is positive.

Proof. (2) implies (1) is obvious. Now, given $n \in \mathbb{N}$, and $a_1, \ldots a_n \in \mathcal{A}$, we consider the matrix $A = (\varphi(a_i^* a_j))_{i,j=1}^n$. For $\xi = (\xi_1, \ldots, \xi_n) \in \mathbb{C}^n$, we have

$$\begin{aligned} \langle A\xi, \xi \rangle = \sum_{i,j=1}^{n} \varphi(a_i^* a_j) \xi_j \overline{\xi_i} &= \varphi(\sum_{i,j=1}^{n} \xi_j \overline{\xi_i} a_i^* a_j) \\ &= \varphi((\sum_{i=1}^{n} \xi_i a_i)^* (\sum_{i=1}^{n} \xi_i a_i)) \geq 0. \end{aligned}$$

$\square$

Now, we start to prove the positivity of φ; that is, Theorem 10. At first, by Lemma 3, for $i(1) \neq \ldots \neq i(n)$, and $j(1) \neq \ldots \neq j(m)$, we have the following observation that

$$(i(1), \ldots, i(n)) \neq (j(1), \ldots, j(m)) \Rightarrow \varphi(a^* b) = 0 \tag{2.4}$$

for all $a \in W_{i(1),\dots,i(n)}$, and $b \in W_{j(1)\dots,j(m)}$. Now, we consider an element $a \in \mathcal{A}$ and write it as

$$a = \sum_{n=0}^{N} \sum_{\substack{i(1),\dots,i(n)\in I;\\ i(1)\neq\dots\neq i(n)}} a_{i(1)\dots i(n)}$$

for some $N \geq 0$ and where $a_{i(1)\dots i(n)} \in W_{i(1),\dots,i(n)}$ for $0 \leq n \leq N$ and $i(1),\dots,i(n) \in I$ are such that $i(1) \neq i(2) \neq \dots \neq i(n)$. Thus,

$$\begin{aligned}
\varphi(a^*a) &= \sum_{m,n=0}^{N} \sum_{\substack{i(1),\dots,i(n)\in I;\\ i(1)\neq\dots\neq i(n)}} \sum_{\substack{j(1),\dots,j(m)\in I;\\ j(1)\neq\dots\neq j(m)}} \varphi(a^*_{i(1),\dots,i(n)} a_{j(1),\dots,j(m)}) \\
&= \sum_{n=0}^{N} \sum_{\substack{i(1),\dots,i(n)\in I;\\ i(1)\neq\dots\neq i(n)}} \varphi(a^*_{i(1),\dots,i(n)} a_{i(1),\dots,i(n)}),
\end{aligned}$$

where at the last equality sign we made use of the implication (2.4).

Therefore, it suffices to show that $\varphi(b^*b) \geq 0$ for $b \in W_{i(1),\dots,i(n)}$. Now, we fix such a b, and write it as $b = \sum_{k=1}^{p} a_1^{(k)} \dots a_n^{(k)}$, where $a_m^{(k)} \in \mathcal{A}^\circ_{i(m)}$ for $1 \leq m \leq n, 1 \leq k \leq p$. Thus, we have

$$\begin{aligned}
\varphi(b^*b) &= \sum_{k,l=1}^{p} \varphi((a_1^{(k)} \dots a_n^{(k)})^* \cdot (a_1^{(l)} \dots a_n^{(l)})) = \sum_{k,l=1}^{p} \varphi(a_n^{(k)*} \dots a_1^{(k)*} \cdot a_1^{(l)} \dots a_n^{(l)}) \\
&= \sum_{k,l=1}^{p} \varphi(a_1^{(k)*} a_1^{(l)}) \dots \varphi(a_n^{(k)*} a_n^{(l)}) \quad \text{by Lemma 3} \\
&= \sum_{k,l=1}^{p} \varphi_{i(1)}(a_1^{(k)*} a_1^{(l)}) \dots \varphi_{i(n)}(a_n^{(k)*} a_n^{(l)}).
\end{aligned}$$

The last equality hold since $\varphi|_{\mathcal{A}_i} = \varphi_i$ for all $i \in I$. Now, for $1 \leq m \leq n$, let the matrix $B_m = (\varphi_{i(m)}(a_m^{(k)*} a_m^{(l)}))_{k,l=1}^p \in M_p(\mathbb{C})$, and let S be the Schur product of $B_1 B_2 \dots B_n$. By Lemma 13, $B_1, \dots B_n$ are positive, and hence by Lemma 12, S is positive. Finally, we note that $\varphi(b^*b) =$ sum of all entries in the Schur product $B_1 \dots B_n = S$; therefore, $\varphi(b^*b) = \langle S\eta, \eta\rangle \geq 0$ where $\eta = (1, 1, \dots, 1) \in \mathbb{C}^p$. Thus, we complete the proof of Theorem 10.

We remark that the free product state φ is faithful if all φ_i are faithful, see [2] for a proof.

2.3 Free Product of C^*-Probability Spaces

We recall that $\mathcal{A}$ is a unital C^*-*algebra* if $\mathcal{A}$ is a unital $*$-algebra which is endowed with a norm $\|\cdot\| : \mathcal{A} \to [0, \infty)$ such that $(\mathcal{A}, \|\cdot\|)$ is a Banach space, and for all $a, b \in \mathcal{A}$, $\|ab\| \leq \|a\| \cdot \|b\|$ and $\|a^*a\| = \|a\|^2$. A C^*-*probability space* is a $*$-probability space $(\mathcal{A}, \varphi)$ where the $*$-algebra $\mathcal{A}$ is required to be a unital C^*-algebra.

In the beginning of this section, we consider the framework of $*$-probability space. For a $*$-probability space $(\mathcal{A}, \varphi)$, we say that (H, π, ξ) is a triple which represent $(\mathcal{A}, \varphi)$ means that H is a Hlibert space, and $\pi : \mathcal{A} \to B(H)$ is a unital $*$-homomorphism and $\xi \in H$ such that $\varphi(a) = \langle \pi(a)\xi, \xi\rangle$ for $a \in \mathcal{A}$.

Let $(\mathcal{A}, \varphi)$ be a $*$-probability space. Consider the positive definite sesquilinear form on $\mathcal{A}$ defined by $\langle a, b \rangle = \varphi(b^*a)$ for all $a, b \in \mathcal{A}$. Let $\mathcal{N} = \{a \in \mathcal{A} | \langle a, a \rangle = 0\}$. Since $|\langle a, b \rangle|^2 = |\varphi(b^*a)|^2 \leq \varphi(a^*a)\varphi(b^*b)$, $\mathcal{N}$ also can be described as $\{a \in \mathcal{A} | \langle a, b \rangle = 0 \text{ for all } b \in \mathcal{A}\}$ and is therefore a vector subspace of $\mathcal{A}$. Now, if we consider $(\mathcal{A}/\mathcal{N}, \langle , \rangle)$, then $(\mathcal{A}/\mathcal{N}, \langle , \rangle)$ is an inner product space. Furthermore, if we take the completion of $(\mathcal{A}/\mathcal{N}, \langle , \rangle)$, it becomes a Hilbert space and we denote it by $L^2(\mathcal{A}, \varphi)$.

Remark 14. *If we consider the linear map from $\mathcal{A}$ into $L^2(\mathcal{A}, \varphi)$ by $a \to \hat{a}$, then $\{\hat{a} | a \in \mathcal{A}\}$ is a dense subspace of $L^2(\mathcal{A}, \varphi)$ and $\langle \hat{a}, \hat{b} \rangle = \varphi(b^*a)$, $a, b \in \mathcal{A}$.*

Proposition 15. *Let $(\mathcal{A}, \varphi)$ be a $*$-probabiity space, and we assume that*

$$\mathcal{A} = span\{u \in \mathcal{A} | u \text{ is a unitary}\}. \tag{2.5}$$

Then for all $a \in \mathcal{A}$, there is a unique $\pi(a) \in B(L^2(\mathcal{A}, \varphi))$ such that $\pi(a)\hat{b} = \widehat{ab}$ for all $b \in \mathcal{A}$. The map $\pi : \mathcal{A} \to B(L^2(\mathcal{A}, \varphi))$ so defined is a unital $*$-homomorphism. Moreover, the triple $(L^2(\mathcal{A}, \varphi), \pi, \hat{1})$ is a representation of $(\mathcal{A}, \varphi)$, where $\hat{1}$ is defined according to the conventions of notation in Remark 14, with $1 = 1_{\mathcal{A}}$, the unit of $\mathcal{A}$.

Proof. Given $a \in \mathcal{A}$. It is obvious that the map $\hat{b} \to \widehat{ab}$ is linear. Now, we claim that the map is bounded. That is, there is $k(a) \geq 0$ such that

$$\|\widehat{ab}\| \leq k(a)\|\hat{b}\|, \quad \text{for all } b \in \mathcal{A}. \tag{2.6}$$

Let a set C be the collection of all $a \in \mathcal{A}$ with the property: there is $k(a) > 0$ such that (2.6) holds. Then, it is easy to see that C is a vector space. On the other hand, if $u \in \mathcal{A}$ is a unitary, then

$$\|\widehat{ub}\| = \langle \widehat{ub}, \widehat{ub} \rangle^{1/2} = \varphi(b^*u^*ub)^{1/2} = \varphi(b^*b)^{1/2} = \|\hat{b}\| \text{ for all } b \in \mathcal{A}.$$

Thus, C contains all the unitaries of $\mathcal{A}$. By our assumption of $\mathcal{A}$, $\mathcal{A} \subseteq C$. Hence, this map is bounded. Since $\{\hat{b} | b \in \mathcal{A}\}$ is dense in $L^2(\mathcal{A}, \varphi)$, there is a unique $\pi(a) \in B(L^2(\mathcal{A}, \varphi))$ such that $\pi(a)\hat{b} = \widehat{ab}$, for $b \in \mathcal{A}$.

We note that for $a \in \mathcal{A}$, $\langle \pi(a)\hat{1}, \hat{1} \rangle = \langle \hat{a}, \hat{1} \rangle = \varphi(1^*a) = \varphi(a)$. Hence, $(L^2(\mathcal{A}), \pi, \hat{1})$ is a representation of $(\mathcal{A}, \varphi)$. $\square$

This representation of $(\mathcal{A}, \varphi)$ described in Proposition 15 is called the *GNS representation*.

Proposition 16. *Let $\mathcal{A}$ is a unital C^*-algebra, then $\mathcal{A} = span\{u \in \mathcal{A} | u \text{ is a unitary }\}$.*

Proof. Note that for each $a \in \mathcal{A}$, a can write as the span of two selfadjoint elements as follows:

$$a = \frac{a + a^*}{2} + i\frac{a - a^*}{2i}.$$

It suffices to prove $a \in \mathcal{A}$ is the span of unitaries under the assumption that $a = a^*$ and $\|a\| \leq 1$. Now we consider $u = a + i\sqrt{1 - a^2}$, then

$$u^*u = (a - i\sqrt{1 - a^2})(a + i\sqrt{1 - a^2}) = 1 = (a + i\sqrt{1 - a^2})(a - i\sqrt{1 - a^2}) = uu^*.$$

Thus, u is a unitary and $a = (u + u^*)/2$. Hence, a is the linear combination of unitaries. $\square$

Remark 17. *By Propisition 16, we know that the the hypothesis (2.5) that $\mathcal{A}$ is the linear span of its unitaries is for instance satisfied whenever $\mathcal{A}$ is a unital C^*-algebra. However, the hypothesis (2.5) cannot be removed if we consider $*$-probability space instead of C^*-probability space. In this case, we may not have the boundedness of the operator $\widehat{b} \to \widehat{ab}$ on $L^2(\mathcal{A}, \varphi)$. For example, let γ be the standard normal distribution on $\mathbb{R}$. Consider*

$$\mathcal{A} = L^{\infty-}(\mathbb{R}, \gamma) = \bigcap_{1 \leq p < \infty} L^p(\mathbb{R}, \gamma),$$

and $\varphi : \mathcal{A} \to \mathbb{C}$ is defined by

$$\varphi(f) = \int_{-\infty}^{\infty} f(t) d\gamma(t), \;\; f \in \mathcal{A}.$$

If we define $f^* = \bar{f}$, $(\mathcal{A}, \varphi)$ becomes a $*$-probability space. It is obvious that

$$\varphi(f^* f) = \int_{\mathbb{R}} |f(t)|^2 d\gamma(t) = 0 \;\Rightarrow\; f = 0 \;\gamma - a.e.$$

Thus, $L^2(\mathcal{A}, \varphi)$ is the completion of the space $(\mathcal{A}, \langle,\rangle_{\mathcal{A}})$. We are going to prove that $L^2(\mathcal{A}, \varphi) = L^2(\mathbb{R}, \gamma)$. At first, we observe that the inner product $\langle,\rangle_{\mathcal{A}}$ on $\mathcal{A}$ induces the same norm as $L^2(\mathbb{R}, \gamma)$. Now, let $f \in L^2(\mathcal{A}, \varphi)$. Case 1: f is from $(\mathcal{A}, \langle,\rangle_{\mathcal{A}})$, then $f \in L^2(\mathbb{R}, \gamma)$ since $f \in \mathcal{A}$. Case 2: f is a limit point of $(\mathcal{A}, \langle,\rangle_{\mathcal{A}})$, then we can find $g \in (\mathcal{A}, \langle,\rangle_{\mathcal{A}})$ such that $\|f - g\|_{L^2(\gamma)} < 1$. Thus, we have

$$\begin{aligned}\int_{\mathbb{R}} |f(t)|^2 d\gamma(t) = \int_{\mathbb{R}} (|f(t) - g(t)| + |g(t)|)^2 d\gamma(t) &\leq 2\int_{\mathbb{R}} |f(t) - g(t)|^2 d\gamma(t) + 2\int_{\mathbb{R}} |g(t)|^2 d\gamma(t) \\ &< 2 + 2\int_{\mathbb{R}} |g(t)|^2 d\gamma(t) < \infty,\end{aligned}$$

which implies that $f \in L^2(\mathbb{R}, \gamma)$. Conversely, given $f \in L^2(\mathbb{R}, \gamma)$, then there is a sequence of simple functions $f_n \in L^2(\mathbb{R}, \gamma)$ such that $\|f_n - f\|_{L^2(\gamma)} \to 0$ as $n \to \infty$. Note that for each $n \in \mathbb{N}$, simple function $f_n \in \mathcal{A}$. Hence, f is a limit point of $\mathcal{A}$, and therefore $f \in L^2(\mathcal{A}, \varphi)$. As a result, $L^2(\mathcal{A}, \varphi) = L^2(\mathbb{R}, \gamma)$.

If we let $f(t) = t$ on $\mathbb{R}$, then $f \in \mathcal{A}$, and for $n \in \mathbb{N}$, consider $g_n = 1/\sqrt{C_n} I_{[n,n+1]}$ where $C_n = \gamma([n, n+1])$. Then $g_n \in L^2(\mathcal{A}, \varphi)$ with $\|g_n\|_{L^2(\gamma)} = 1$, and we have

$$\|fg_n\|^2_{L^2(\gamma)} = \int_{\mathbb{R}} |f(t)g_n(t)|^2 d\gamma(t) = \frac{1}{C_n}\int_n^{n+1} t^2 d\gamma(t) \geq \frac{1}{C_n}\int_n^{n+1} n^2 d\gamma(t) = n^2 \to \infty \text{ as } n \to \infty.$$

Thus, the map $g \to fg$ is not a bounded operator on $L^2(\mathcal{A}, \varphi)$. Therefore, $\mathcal{A}$ doesn't have the property (2.5); as a result, $\mathcal{A}$ is not a C^*-algebra with respect to any given norm. Alternatively, we can show directly that $\mathcal{A} \neq span\{u \in \mathcal{A} | u \text{ is a unitary }\}$ as follows: suppose that $\mathcal{A}$ is the linear span of its unitaries. Note that if u is a unitary in $\mathcal{A}$, $u(t)u(t)^* = |u(t)|^2 = 1 \;\gamma-a.e.$ Let $f(t) = t$. Under our assumption, there are finite complex numbers $\alpha_1, \ldots, \alpha_n$ and unitaries $u_1, \ldots, u_n$ such that

$$f(t) = \sum_{j=1}^n \alpha_j u_j(t).$$

Then we have

$$|f(t)| \leq \sum_{j=1}^n |\alpha_j||u_j(t)| = \sum_{j=1}^n |\alpha_j| < \infty \;\gamma - a.e,$$

which contradicts to the fact that f is not bounded γ-a.e.

Definition 18. Let $\mathcal{A}$ be a unital $*$-algebra. Let H be a Hilbert space and $\pi : \mathcal{A} \to B(H)$ be a unital $*$-homomorphism. A vector $\eta \in H$ is said to be *cyclic* for π if $\overline{\{\pi(a)\eta | a \in \mathcal{A}\}} = H$.

We note that $\{\pi(a)\hat{1} | a \in \mathcal{A}\} = \{\widehat{a1} | a \in \mathcal{A}\} = \{\hat{a} | a \in \mathcal{A}\}$, and hence $\hat{1}$ is cyclic for GNS representation $\pi : \mathcal{A} \to B(L^2(\mathcal{A}, \varphi))$.

Proposition 19. *Let $(\mathcal{A}, \varphi)$ be a $*$-probability space, and assume that $(\mathcal{A}, \varphi)$ satisfies the hypothesis of Proposition 15. Let (H, ρ, ξ) be a representation of $(\mathcal{A}, \varphi)$ such that ξ is cyclic for ρ. Then (H, ρ, ξ) is unitarily equivalent to the GNS representation $(L^2(\mathcal{A}, \varphi), \pi, \hat{1})$. That is, there is a linear operator $U : L^2(\mathcal{A}, \varphi) \to H$ which is bijective and norm preserving such that $U(\hat{1}) = \xi$ and $U\pi(a)U^* = \rho(a)$ for all $a \in \mathcal{A}$.*

Proof. At first, for $a \in \mathcal{A}$, we observe that

$$\|\rho(a)\xi\|^2 = \langle \rho(a)\xi, \rho(a)\xi \rangle = \langle \rho(a)^*\rho(a)\xi, \xi \rangle = \langle \rho(a^*a)\xi, \xi \rangle = \varphi(a^*a) = \langle \hat{a}, \hat{a} \rangle = \|\hat{a}\|^2.$$

Thus, we have $\|\rho(a)\xi\| = \|a\|$. Now, we define $U_0 : \{\hat{a} | a \in \mathcal{A}\} \to H$ by $U_0(\hat{a}) = \rho(a)\xi$. If $\hat{a} = \hat{b}$, then we have

$$0 = \|\hat{a} - \hat{b}\| = \|\widehat{a-b}\| = \|\rho(a-b)\xi\| = \|\rho(a)\xi - \rho(b)\xi\| \Rightarrow \rho(a)\xi = \rho(b)\xi.$$

Thus, U_0 is well-defined, and it is linear and isometric obviously. There is a continuous extension U of U_0 such that $U : L^2(\mathcal{A}, \varphi) \to H$ is linear norm-preserving with $U|_{\{\hat{a}|a\in\mathcal{A}\}} = U_0$. Since U is isometric and $L^2(\mathcal{A}, \varphi)$ is complete, we have $range(U)$ is also complete, hence it is closed in H. On the other hand, $H = \overline{range(U_0)} = \overline{\{\rho(a)\xi | a \in \mathcal{A}\}} \subseteq range(U)$ where the first equality hold since ξ is cyclic for ρ. Therefore, $H = range(U)$. Finally, $U(\hat{1}) = \rho(1)\xi = \xi$ and for $b \in \mathcal{A}$, we have $U\pi(a)\hat{b} = U\widehat{ab} = \rho(ab)\xi = \rho(a)\rho(b)\xi = \rho(a)U\hat{b}$. Hence, $U\pi(a) = \rho(a)U$ which implies $U\pi(a)U^* = \rho(a)$. $\square$

Let $\mathcal{A}$ be a unital $*$-algebra, and let $\pi : \mathcal{A} \to B(H)$ be a unital $*$-homomorphism, an element $\eta \in H$ is called *separating* for π if the map $a \longmapsto \pi(a)\eta$ is one-to-one.

Proposition 20. *Let $(\mathcal{A}, \varphi)$ be a $*$-probability space, which satisfies the hypothesis of Proposition 15. Then φ is faithful if and only if $\hat{1}$ is a separating for π.*

Proof. Suppose that $\hat{1}$ is separating for π. Given $\varphi(a^*a) = 0$. Then

$$0 = \langle \pi(a)\hat{1}, \pi(a)\hat{1} \rangle = \|\pi(a)\hat{1}\|^2 \Rightarrow \pi(a)\hat{1} = 0 \Rightarrow a = 0.$$

Thus, φ is faithful. Conversely, if φ is faithful, we assume that $\pi(a)\hat{1} = 0$, and hence $\hat{a} = 0$. Then, we have

$$\varphi(a^*a) = \langle \pi(a^*a)\hat{1}, \hat{1} \rangle = \langle \pi(a)\hat{1}, \pi(a)\hat{1} \rangle = \langle \hat{a}, \hat{a} \rangle = \|\hat{a}\|^2 = 0.$$

Therefore, φ is faithful implies $a = 0$. $\square$

According to the Proposition 20, we obtain a result: if φ is faithful, then $\pi : \mathcal{A} \to B(L^2(\mathcal{A}, \varphi))$ is one-to-one.

Lemma 21. *Let $(\mathcal{A}_0, \varphi_0)$ be a $*$-probability space such that φ_0 is a faithful trace. Suppose that $\mathcal{A}_0$ satisfies the hypothesis of Proposition 15, and consider the GNS representation $(L^2(\mathcal{A}_0, \varphi_0), \pi, \hat{1})$, as described in that proposition. Let $\mathcal{A} = \overline{\pi(\mathcal{A}_0)} \subseteq B(L^2(\mathcal{A}_0, \varphi_0))$. If $T \in \mathcal{A}$ is such that $T\hat{1} = 0$, then $T = 0$.*

Proof. For $a, b, c \in \mathcal{A}_0$, we have

$$\langle \pi(c)\hat{a}, \hat{b}\rangle = \langle \widehat{ca}, \hat{b}\rangle = \varphi(b^*ca) = \varphi(ab^*c) = \langle \pi(c)\hat{1}, \widehat{ba^*}\rangle.$$

For $T \in \mathcal{A}$, we can approximating T by $\pi(c),\ c \in \mathcal{A}_0$, so $\langle T\hat{a}, \hat{b}\rangle = \langle T\hat{1}, \widehat{ba^*}\rangle$ for all $a, b \in \mathcal{A}_0$. Now, let $T \in \mathcal{A}$ such that $T\hat{1} = 0$, then $\langle T\hat{a}, \hat{b}\rangle = 0$ for all $a, b \in \mathcal{A}_0$, which implies $T = 0$ on $\{\hat{a} | a \in \mathcal{A}_0\}$. Since $\{\hat{a} | a \in \mathcal{A}_0\}$ is dense on $L^2(\mathcal{A}_0, \varphi_0)$ and T is a bouned linear operator, we obtain $T = 0$. □

Theorem 22. *Let $(\mathcal{A}_i, \varphi_i)_{i\in I}$ be C^*-probability spaces such that the functionals $\varphi_i : \mathcal{A}_i \to \mathbb{C}, i \in I$, are faithful traces. Then there is a C^*-probability space $(\mathcal{A}, \varphi)$ with φ a faithful trace, and a family of norm-preserving unital $*$-homomorphisms $W_i : \mathcal{A}_i \to \mathcal{A},\ i \in I$, such that:*
(1) $\varphi \circ W_i = \varphi_i$, for $i \in I$;
(2) the unital C^*-subalgebras $(W_i(\mathcal{A}_i))_{i\in I}$ form a free family in $(\mathcal{A}, \varphi)$.

Proof. In order to construct $(\mathcal{A}, \varphi)$, let us first consider the free product of $*$-probability spaces $(\mathcal{A}_0, \varphi_0) = *_{i\in I}(\mathcal{A}_i, \varphi_i)$. Then, φ_0 is faithful trace on $\mathcal{A}_0$. Now, we set $W = span\{u \in \mathcal{A}_0 | u$ is unitary$\}$ and we claim that W is exactly $\mathcal{A}_0$. For each $i \in I$,

$$W \supset span\{u \in \mathcal{A}_i | u \text{ is unitary}\} = \mathcal{A}_i \ \Rightarrow W \supset \bigcup_{i\in I} \mathcal{A}_i.$$

On the other hand, W is a $*$-subalgebra of $\mathcal{A}_0$, which implies W contains the unital $*$-algebra of $\mathcal{A}_0$ generated by $\bigcup_{i\in I} \mathcal{A}_i$, which is all of $\mathcal{A}_0$.

Since $(\mathcal{A}_0, \varphi_0)$ satisfies the hypothesis of Proposition 15, we consider the GNS representation $(L^2(\mathcal{A}_0, \varphi_0), \pi, \hat{1})$ for $(\mathcal{A}_0, \varphi_0)$. Since φ_0 is faithful, we have $\pi : \mathcal{A}_0 \to B(L^2(\mathcal{A}_0, \varphi_0))$ is one-to-one.

Now, we consider the unital C^*-probability space $\mathcal{A} = \overline{\pi(\mathcal{A}_0)} \subseteq B(L^2(\mathcal{A}_0, \varphi_0))$. For each $i \in I$, we note that $\mathcal{A}_i \subset \mathcal{A}_0 = domain(\pi)$, and $\mathcal{A}$ contains the range of $\mathcal{A}_i$, we can define a unital $*$-homomorphism map $W_i : \mathcal{A}_i \to \mathcal{A}$ which is a restriction of π. Since π is one-to-one, we have W_i is also one-to-one. Note that $\mathcal{A}_i, \mathcal{A}$ are C^*-algebras, and W_i is $*$-homomorphism one-to-one, which implies that W_i is norm preserving.

Let $\varphi : \mathcal{A} \to \mathbb{C}$ be the positive linear functional defined by $\varphi(T) = \langle T\hat{1}, \hat{1}\rangle$ for $T \in \mathcal{A}$. Then, $(\mathcal{A}, \varphi)$ is a C^*-probability space clearly. For a given $T \in \mathcal{A}$ such that $\varphi(T^*T) = 0$, we have $0 = \varphi(T^*T) = \|T\hat{1}\|^2$, which implies $T\hat{1} = 0 \Rightarrow T = 0$. Thus φ is faithful.

We observe that $\varphi(\pi(a)) = \langle \pi(a)\hat{1}, \hat{1}\rangle = \langle \hat{a}, \hat{1}\rangle = \varphi_0(a)$ for all $a \in \mathcal{A}_0$. In particular, $\varphi \circ W_i = \varphi_i$ for all $i \in I$. We check φ is a trace. Given $a, b \in \mathcal{A}$, if $a, b \in \pi(\mathcal{A}_0)$, then there exists unique $a', b' \in \mathcal{A}_0$ such that $\pi(a') = a,\ \pi(b') = b$ since π is one-to-one. Thus,

$$\varphi(ab) = \varphi \circ \pi(a'b') = \varphi_0(a'b') = \varphi_0(b'a') = \varphi \circ \pi(b'a') = \varphi(ba).$$

Now, if $a, b \in \mathcal{A}$, then there are two sequences $\{a_n\}_{n=0}^{\infty}, \{b_n\}_{n=0}^{\infty}$ such that $a_n \to a,\ b_n \to b$ as $n \to \infty$. Then, $\varphi(ab) = \lim_{n\to\infty} \varphi(a_n b_n) = \lim_{n\to\infty} \varphi(b_n a_n) = \varphi(ba)$.

Finally, we check that $\{W_i(\mathcal{A}_i)\}_{i\in I}$ are free in $(\mathcal{A}, \varphi)$. For $n \in \mathbb{N}$, given $i(1) \neq \ldots \neq i(n)$ in I, and $a_j \in W_{i(j)}(\mathcal{A}_{i(j)})$ with $\varphi(a_j) = 0$ where $j = 1, \ldots, n$. Since $W_{i(j)}$ is one-to-one for each $j = 1, \ldots, n$, there is a unique $b_j \in \mathcal{A}_{i(j)}$ such that $W_{i(j)}(b_j) = a_j$. Thus, we have

$$\begin{aligned} \varphi(a_1 a_2 \ldots a_n) &= \varphi(W_{i(1)}(b_1) \ldots W_{i(n)}(b_n)) = \varphi(\pi(b_1) \ldots \pi(b_n)) = \varphi(\pi(b_1 \ldots b_n)) \\ &= \varphi \circ \pi(b_1 \ldots b_n) = \varphi_0(b_1 \ldots b_n) = 0. \end{aligned}$$

The last equality hold since $\varphi_0(b_j) = \varphi(W_{i(j)}(b_j)) = \varphi(a_j) = 0$ and the fact that $(\mathcal{A}_i)_{i\in I}$ are free in $(\mathcal{A}_0, \varphi_0)$. $\square$

2.4 The Cauchy Transform

Notation 23. $\mathbb{C}^+ = \{z \in \mathbb{C} | Im(z) > 0\}$:*the complex upper half plane.*
$\mathbb{C}^- = \{z \in \mathbb{C} | Im(z) < 0\}$:*the complex lower half plane.*

Definition 24. Let ν be a probability measure on $\mathbb{R}$, and for $z \in \mathbb{C}^+$, the *Cauchy transform* of ν is defined by

$$G_\nu(z) = \int_{\mathbb{R}} \frac{1}{z-t} d\nu(t).$$

We note that for $z \in \mathbb{C}^+,\ t \in \mathbb{R}$, $|z - t| \geq Im(z) \Rightarrow |z - t|^{-1} \leq Im(z)^{-1}$. Then,

$$\left|\int_{\mathbb{R}} \frac{1}{z-t} d\nu(t)\right| \leq \int_{\mathbb{R}} \left|\frac{1}{z-t}\right| d\nu(t) \leq \int_{\mathbb{R}} \frac{1}{Im(z)} d\nu(t) = \frac{1}{Im(z)} < \infty.$$

Thus, the Cauchy transform G_ν is well-defined.

Proposition 25. *G_ν is an analytic function on $\mathbb{C}^+$ with range contained in $\mathbb{C}^-$.*

Proof. Fixed $z \in \mathbb{C}^+$, we let $w \in \mathbb{C}$ such that $Im(w) \neq 0$ and $|z - w| < \frac{1}{2}|Im(w)|$. Then for $t \in \mathbb{R}$,

$$\left|\frac{z-w}{t-w}\right| < \frac{|Im(w)|}{2} \frac{1}{|Im(w)|} = \frac{1}{2}.$$

Therefore,

$$\sum_{n=0}^{\infty} \left(\frac{z-w}{t-w}\right)^n \text{ converges uniformly to } \frac{t-w}{t-z} \Rightarrow \frac{1}{z-t} = -\sum_{n=0}^{\infty} \frac{(z-w)^n}{(t-w)^{(n+1)}} \text{ on } |z-w| < \frac{1}{2}|Im(w)|.$$

Hence,

$$G_\nu(z) = -\sum_{n=0}^{\infty} \left[\int_{\mathbb{R}} (t-w)^{-(n+1)} d\nu(t)\right](z-w)^n \text{ is analytic on } |z-w| < \frac{1}{2}|Im(w)|.$$

Finally, since $z \in \mathbb{C}^+,\ Im(z) > 0$, we note that $Im((z-t)^{-1}) < 0$ for $t \in \mathbb{R}$, and then $Im(G_\nu(z)) < 0$. Thus, $G_\nu(\mathbb{C}^+) \subseteq C^-$. $\square$

Proposition 26. *Let G_ν be the Cauchy transform of a probability measure ν on $\mathbb{R}$. Then*

$$\lim_{y\to\infty} iyG_\nu(iy) = 1 \text{ and } \sup_{y>0, x\in\mathbb{R}} y|G_\nu(x+iy)| = 1.$$

Proof. Note that

$$Re(G_\nu(iy)) = \int \frac{-t}{y^2+t^2} d\nu(t), \text{ and } Im(G_\nu(iy)) = \int \frac{-y}{y^2+t^2} d\nu(t).$$

Since $|(-y^2)/(y^2+t^2)| = |-(1+(t/y)^2)^{-1}| \leq 1$ for $y > 0$, and $-(1+(t/y)^2)^{-1} \to -1$ as $y \to \infty$, by Dominated Convergence Theorem we have

$$yIm(G_\nu(iy)) = \int_{\mathbb{R}} \frac{-y^2}{y^2+t^2} d\nu(t) = -\int_{\mathbb{R}} \frac{1}{1+(t/y)^2} d\nu(t) \ \to -\int_{\mathbb{R}} d\nu(t) = -1 \text{ as } y \to \infty.$$

On the other hand, for $y > 0,\ t \in \mathbb{R},\ |yt/(y^2+t^2)| \leq 1/2$, and $yt/(y^2+t^2)$ converges to 0 as $y \to \infty$. By Dominated Convergence Theorem again, we have

$$yRe(G_\nu(iy)) = \int_{\mathbb{R}} \frac{-yt}{y^2+t^2} d\nu(t) \to \int_{\mathbb{R}} 0 d\nu(t) = 0 \text{ as } y \to \infty.$$

Thus, we have proved the first equality.

Now, we observe that for $y > 0$ and $z = x+iy$,

$$y|G_\nu(z)| \leq \int_{\mathbb{R}} \frac{y}{|z-t|} d\nu(t) = \int_{\mathbb{R}} \frac{y}{\sqrt{(x-t)^2+y^2}} d\nu(t) \leq 1.$$

Thus, $\sup_{y\geq 0, x\in\mathbb{R}} y|G(x+iy)| \leq 1$. By the first part, the supremum is 1. ☐

Before we start next theorem, we recall a classical result of R. Nevanlinna: suppose that $\varphi : \mathbb{C}^+ \to \mathbb{C}^+$ is analytic, then there is a unique finite positive Borel measure σ on $\mathbb{R}$ and real numbers α and β, with $\beta \geq 0$ such that for $z \in \mathbb{C}^+$,

$$\varphi(z) = \alpha + \beta z + \int_{\mathbb{R}} \frac{1+tz}{t-z} d\sigma(t). \tag{2.7}$$

The equation (2.7) is called the *Nevanlinna representation* of F.

Theorem 27. *Suppose $G : \mathbb{C}^+ \to \mathbb{C}^-$ is analytic and $\limsup_{y\to\infty} y|G(iy)| = c < \infty$. Then there is a unique positive Borel measure ν on $\mathbb{R}$ such that*

$$G(z) = \int_{\mathbb{R}} \frac{1}{z-t} d\nu(t) \text{ and } \nu(\mathbb{R}) = c.$$

Proof. Note that $-G : \mathbb{C}^+ \to \mathbb{C}^+$ is analytic, then there is a unique finite positive measure σ on $\mathbb{R}$ such that $G(z) = \alpha + \beta z + \int_{\mathbb{R}} (1+tz)/(z-t) d\sigma(t)$ with $\alpha \in \mathbb{R}$ and $\beta \leq 0$. First, consider $Re(iyG(iy))$, we obtain that

$$2c \geq Re(iyG(iy)) = y^2(-\beta + \int_{\mathbb{R}} \frac{1+t^2}{y^2+t^2} d\sigma(t)) \geq -\beta y^2,$$

for all $y > 0$ large enough. This implies $\beta = 0$. Therefore, we have

$$y^2 \int_{\mathbb{R}} \frac{1+t^2}{y^2+t^2} d\sigma(t) = \int_{\mathbb{R}} \frac{1+t^2}{1+(t/y)^2} d\sigma(t) \leq 2c,$$

for all $y > 0$ sufficiently large. Thus, by Monotone Convergence Theorem, $\int_{\mathbb{R}}(1+t^2)d\sigma(t) \le 2c$ and hence σ has the second moment. From the imaginary part of $iyG(iy)$ we get that for $y > 0$ sufficiently large, we have

$$y|\alpha + \int_{\mathbb{R}} \frac{t(y^2-1)}{t^2+y^2} d\sigma(t)| \le 2c,$$

which implies that $\alpha = -\lim_{y\to\infty} \int_{\mathbb{R}} t(y^2-1)/(t^2+y^2)d\sigma(t)$.

Since $|(y^2-1)/(t^2+y^2)| \le 1$ for $y \ge 1$ and σ has the second moment (and hence the first moment), we apply the Dominated Convergence Theorem and conclude that

$$\alpha = -\lim_{y\to\infty} t\frac{1-y^{-2}}{1+(t/y)^2} d\sigma(t) = -\int_{\mathbb{R}} t d\sigma(t).$$

Hence, if we let $\nu(E) = \int_E (1+t^2)d\sigma(t)$, then ν is a finite measure since σ has the second moment.

$$G(z) = \int_{\mathbb{R}} (-t + \frac{1+tz}{z-t}) d\sigma(t) = \int_{\mathbb{R}} \frac{1}{z-t}(1+t^2)d\sigma(t) = \int_{\mathbb{R}} \frac{1}{z-t} d\nu(t).$$

So, G is the Cauchy transform of the positive Borel measure ν. By Proposition 26, the imaginary part of $iyG(iy)$ goes to 0 and the real part is positive. As a result, we have

$$c = \limsup_{y\to\infty} |iyG(iy)| = \lim_{y\to\infty} Re(iyG(iy)) = \int_{\mathbb{R}} (1+t^2)d\sigma(t) = \nu(\mathbb{R}).$$

$\square$

Theorem 28. *Suppose that ν is a probability measure on $\mathbb{R}$ and G is its Cauchy transform. For $a < b$, we have*

$$-\lim_{y\to 0^+} \frac{1}{\pi} \int_a^b Im(G_\nu(x+iy))dx = \nu((a,b)) + \frac{1}{2}\nu(\{a,b\}). \tag{2.8}$$

If ν_1 and ν_2 are probability measures with $G_{\nu_1} = G_{\nu_2}$, then $\nu_1 = \nu_2$.

Proof. Note that

$$Im(G_\nu(x+iy)) = \int_{\mathbb{R}} Im(\frac{1}{x-t+iy}) d\nu(t) = \int_{\mathbb{R}} \frac{-y}{(x-t)^2+y^2} d\nu(t).$$

Thus,

$$\begin{aligned}\int_a^b Im(G_\nu(x+iy))dx &= \int_{\mathbb{R}}\int_a^b \frac{-y}{(x-t)^2+y^2} dx d\nu(t) = -\int_{\mathbb{R}}\int_{(a-t)/y}^{(b-t)/y} \frac{1}{1+\tilde{x}^2} d\tilde{x} d\nu(t) \\ &= -\int_{\mathbb{R}} [\tan^{-1}(\frac{b-t}{y}) - \tan^{-1}(\frac{a-t}{y})] d\nu(t),\end{aligned}$$

where $\tilde{x} = (x-t)/y$.

Let $f(y,t) = \tan^{-1}((b-t)/y) - \tan^{-1}((a-t)/y)$ and

$$f(t) = \begin{cases} 0 & \text{if } t \notin [a,b] \\ \pi/2 & \text{if } t \in \{a,b\} \\ \pi & \text{if } t \in (a,b) \end{cases}.$$

Then, $\lim_{y\to 0^+} f(y,t) = f(t)$, and for $y > 0$ and for all t, we have $|f(y,t)| \leq \pi$. By Dominated Convergence Theorem,

$$\begin{aligned}\lim_{y\to 0^+}\int_a^b Im(G_\nu(x+iy))dx &= -\lim_{y\to 0^+}\int_{\mathbb{R}} f(y,t)d\nu(t) = -\int_{\mathbb{R}} f(t)d\nu(t)\\ &= -\pi(\nu((a,b)) + \frac{1}{2}\nu(\{a,b\})).\end{aligned}$$

Now, assume that $G_{\nu_1} = G_{\nu_2}$. By the equation (2.8), we have $\nu_1((a,b)) = \nu_2((a,b))$ for all a and b are not atoms of ν_1 and not atoms of ν_2. Let S be the collection of atoms of ν_1 and ν_2, we note that S is countable. Given an interval (a,b). Then, we construct a decreasing sequence $\{a_n\}_{n=1}^{\infty} \in (a,b)\backslash S$ such that $a_n \to a$ as $n \to \infty$ as follows:

Since $(a,b)\backslash S$ is dense in (a,b), there is a_1 in $(a,b)\backslash S$ such that $a_1 < \min\{a+1, b\}$. Suppose that we already have $a_1, \ldots, a_{n-1}$, where $a < a_k < \min\{a+1/k, b\}$ for $1 \leq k \leq n-1$. Since $(a, a_{n-1})\backslash S$ is dense in (a, a_{n-1}), there is $a_n \in (a, a_{n-1})\backslash S$ such that $a_n < \min\{a+1/n, b\}$. Thus, by induction, we can define a decreasing sequence $\{a_n\}_{n=0}^{\infty}$ in $(a,b)\backslash S$ such that $a < a_n < a+1/n$ for all $n \in \mathbb{N}$, and hence $a_n \to a$ as $n \to \infty$.

By using similar way, we can construct a increasing sequence $\{b_n\}_{n=1}^{\infty}$ in $(a,b)\backslash S$ such that $b_n > a_n$ and $b_n \to b$ as $n \to \infty$. Thus, $(a,b) = \cup_{n=1}^{\infty}(a_n, b_n)$ where a_n and b_n are not in S. Then, we get

$$\nu_1((a,b)) = \lim_{n\to\infty}\nu_1((a_n,b_n)) = \lim_{n\to\infty}\nu_2((a_n,b_n)) = \nu_2((a,b)).$$

This shows that ν_1 and ν_2 agree on all open intervals and thus are equal. $\square$

Theorem 29. *Let ν be a probability measure on $\mathbb{R}$ with support contained in the interval $[-r,r]$ and let G_ν be its Cauchy transform. Then*
(1) G_ν is one-to-one on $\{z||z| > 4r\}$;
(2) $\{z|0 < |z| < 1/6r\} \subseteq \{G_\nu(z)||z| > 4r\}$;
(3) there is an analytic function R_ν on $\{z||z| < 1/(6r)\}$ such that $G_\nu(R_\nu(z) + 1/z) = z$ for $0 < |z| < 1/(6r)$.

Proof. (1) For $n \in \mathbb{N}$, let α_n be the n-th moment of ν, and let $\alpha_0 = 1$. Then $|\alpha_n| \leq \int_{\mathbb{R}} |t|^n d\nu(t) \leq r^n$. Let

$$f(z) = G_\nu(\frac{1}{z}) = \int_{\mathbb{R}} \frac{1}{1/z - t}d\nu(t) = z\int_{\mathbb{R}} \frac{1}{1-tz}d\nu(t).$$

For $|z| < 1/r$ and $t \in supp(\nu)$, we have $|zt| < 1$, then $\sum_{n=0}^{\infty}(zt)^n$ converges uniformly on $supp(\nu)$. On compact subsets of $\{z||z| < 1/r\}$,

$$\begin{aligned}\sum_{n=0}^{\infty}\alpha_n z^{n+1} &= z\sum_{n=0}^{\infty}\alpha_n z^n = z\sum_{n=0}^{\infty}\int_{\mathbb{R}} t^n z^n d\nu(t) = z\int_{\mathbb{R}}\sum_{n=0}^{\infty} t^n z^n d\nu(t)\\ &= z\int_{\mathbb{R}}\frac{1}{1-tz}d\nu(t) = f(z).\end{aligned}$$

So $\sum_{n=0}^{\infty}\alpha_n z^{n+1}$ converges uniformly to $f(z)$ on compact subsets of $\{z||z| < 1/r\}$. Hence, $\sum_{n=0}^{\infty}\alpha_n z^{-(n+1)}$ converges uniformly to $G_\nu(z)$ on compact subsets of $\{z||z| > r\}$.

Now, suppose $z_1, z_2 \in \mathbb{C},\ |z_1|, |z_2| < 1/r$, then

$$\begin{aligned} \left|\frac{f(z_1)-f(z_2)}{z_1-z_2}\right| &\geq Re\left(\frac{f(z_1)-f(z_2)}{z_1-z_2}\right) = Re\left(\int_0^1 \frac{d}{dt}\left[\frac{f(z_1+t(z_2-z_1))}{z_2-z_1}\right]dt\right) \\ &= \int_0^1 Re(f'(z_1+t(z_2-z_1)))dt. \end{aligned}$$

Since

$$Re(-z^n\alpha_n) = Re(-z^n)Re(\alpha_n) \leq |-z|^n r^n = |z|^n r^n \Rightarrow Re(z^n\alpha_n) \geq -|z|^n r^n,$$

we have $Re(-z^n\alpha_n) \geq -|z|^n r^n$.

We note that $\sum\limits_{n=0}^{\infty}(zr)^n$ converges uniformly, then

$$\begin{aligned} Re(f'(z)) &= Re(1+2z\alpha_1+3z^2\alpha_2+\ldots) \geq 1-2|z|r-3|z|^2r^2-\ldots \\ &= 2-(1+2|z|r+3(|z|r)^2+\ldots) = 2-\frac{1}{(1-|z|r)^2}. \end{aligned}$$

Now, for $z \in \mathbb{C}$ such that $|z| < 1/4r,\ Re(f'(z)) \geq 2-1/(1-|z|r)^2 \geq 2-1/(1-1/4)^2 = 2/9$. Hence, for $|z_1|, |z_2| < 1/(4r)$,

$$\left|\frac{f(z_1)-f(z_2)}{z_1-z_2}\right| \geq \int_0^1 Re(f'(z_1+t(z_2-z_1)))dt \geq \int_0^1 \frac{2}{9}dt = \frac{2}{9} \Rightarrow |f(z_1)-f(z_2)| \geq \frac{2}{9}|z_1-z_2|.$$

Therefore, f is one-to-one on $\{z||z| < 1/(4r)\} \iff G_\nu$ is one-to-one on $\{z||z| > 4r\}$.

(2) For any curve Γ in $\mathbb{C},\ w \notin \Gamma$, we let $Ind_\Gamma(w) = (2\pi i)^{-1}\int_\Gamma dz/(z-w)$: the winding number of Γ around w. Let $\Gamma = \{z||z| = 1/(4r)\}$, and then $f(\Gamma) = \{f(z)|z \in \Gamma\}$. Since $f(0) = 0$ and f is one-to-one on $\{z||z| < 1/(4r)\}$, by the argument principle,

$$Ind_{f(\Gamma)}(0) = \frac{1}{2\pi i}\int_\Gamma \frac{f'(z)}{f(z)}dz = 1.$$

Also for $z \in \mathbb{C}$ and $|z| < 1/(4r)$,

$$\begin{aligned} |f(z)| = |z||1+\alpha_1 z+\alpha_2 z^2+\ldots| &\geq |z|(2-(1+r|z|+r^2|z|^2+\ldots)) = |z|(2-\frac{1}{1-r|z|}) \\ &\geq |z|(2-\frac{1}{1-1/4}) = \frac{2}{3}|z|. \end{aligned}$$

If $z \in \mathbb{C}, |z| = 1/(4r)$ and $|f(z)| \geq 1/(6r)$, then $f(\Gamma)$ lies outside the circle $|z| = 1/(6r)$. Thus, for $w \in \{z||z| < 1/(6r)\},\ Ind_{f(\Gamma)}(w) = Ind_{f(\Gamma)(0)} = 1$ since the index is constant on connected components of the complement of $f(\Gamma)$. Hence,

$$1 = Ind_{f(\Gamma)}(w) = \frac{1}{2\pi i}\int_\Gamma \frac{f'(z)}{f(z)-w}dz.$$

By the argyment principle, there is a unique $z \in \mathbb{C},\ |z| < 1/(4r)$ such that $f(z) = w$. Therefore,

$$\{z||z| < \frac{1}{6r}\} \subseteq \{f(z)||z| < \frac{1}{4r}\} \Rightarrow \{z|0 < |z| < \frac{1}{6r}\} \subseteq \{G_\nu(z)||z| < \frac{1}{6r}\}.$$

(3) Consider the inverse function f^{-1} on $\{z||z| < 1/(6r)\}$. Since $f^{-1}(0) = 0$ and $f^{-1}(0) = 1/(f'(0)) = 1 \neq 0$, f^{-1} has a simple zero at 0. We define $K(z) = 1/(f^{-1}(z))$ on $\{z||z| < 1/(6r)\}$. Then, K has a simple

pole at 0. Note that $f^{-1}(z) = zg(z)$ where $g(0) \neq 0$ and then $(f^{-1})(0) = g(0) = 1$; thus,

$$Residue(K, 0) = \lim_{z \to 0} zK(z) = \lim_{z \to 0} \frac{z}{f^{-1}(z)} = \frac{z}{zg(z)} = \frac{1}{g(0)} = 1.$$

Hence, $R_\nu(z) \equiv K(z) - 1/z$ is holomophic on $\{z||z| < 1/(6r)\}$. Now, for $z \in \mathbb{C}, |z| < 1/(6r)$,

$$G_\nu(R(z) + \frac{1}{z}) = G(K(z)) = f(\frac{1}{K(z)}) = f(f^{-1}(z)) = z.$$

$\square$

Chapter 3

Voiculescu's R-Transform Addition Formula

In this chapter, we present a proof of the R-transform addition formula

$$R_{a+b} = R_a + R_b$$

for free random variables a and b. We follow Haagerup's approach [1]. Our main goal (Theorem 34) is to show the existence and uniqueness of the free convolution $\mu \boxplus \nu$ for two compactly supported probability measures μ and ν.

Let $(\mathcal{A}, \varphi)$ be a non-commutative probability space. Given $a \in \mathcal{A}$, *the distribution* μ_a of the element $a \in \mathcal{A}$ is defined as the functional on the space $\mathbb{C}[X]$ by

$$\mu_a(p) = \varphi(p(a)), \quad p \in \mathbb{C}[X]$$

where $\mathbb{C}[X]$ is the space of complex polynomials in one variable. Note that the distribution μ_a is completely determined by the sequence of moments $\{\varphi(a^n)\}_{n=1}^{\infty}$. For a fixed $a \in \mathcal{A}$, we set

$$G_a(\lambda) = \sum_{n=0}^{\infty} \frac{\varphi(a^n)}{\lambda^{n+1}}$$

as a formal Laurent series in λ, and we can compute the inverse function of G as a formal Laurent series

$$G_a^{-1}(z) = \frac{1}{z} + \sum_{n=0}^{\infty} \alpha_n z^n,$$

and the *R-transform* of a is defined by

$$R_a(z) = G_a^{-1}(z) - \frac{1}{z}$$

as a formal series in z. We observe that R-transfom satisfies $G_a(R_a(z) + 1/z) = z$.

Theorem 30. *(R-Transform Addition Formula) If $a, b \in \mathcal{A}$ are free, then $R_{a+b}(z) = R_a(z) + R_b(z)$.*

In order to prove this theorem, we introduce the idea of the full Fock space. At first, we consider the 2-dimension Hilbert space $H = \mathbb{C}^2$ with orthonormal basis $\{e_1, e_2\}$ where $e_1 = (1, 0)$ and $e_2 = (0, 1)$. For each $n \in \mathbb{N}$, $H^{\otimes n}$ is an inner product space where the inner product $\langle\ ,\ \rangle_{H^{\otimes n}} : H^{\otimes n} \times H^{\otimes n} \to \mathbb{C}$ is defined by

$$\langle h_1 \otimes \ldots \otimes h_n, g_1 \otimes \ldots \otimes g_n \rangle_{H^{\otimes n}} = \prod_{i=1}^{n} \langle h_i, g_i \rangle_H,$$

for $h_1, \ldots, h_n, g_1, \ldots, g_n \in H$. Since H is finite dimensional, so is $H^{\otimes n}$. As a result, $H^{\otimes n}$ is a finite dimensional inner product space, which implies that $H^{\otimes n}$ is a Hilbert space. Moreover, due to the fact that $span\{e_{i_1} \otimes \ldots \otimes e_{i_n} \mid i_j = 1 \text{ or } 2,\ \forall j = 1, 2, \ldots, n\}$ is a dense subspace of $H^{\otimes n}$ and $dim(H^{\otimes n}) < \infty$, we have

$$span\{e_{i_1} \otimes \ldots \otimes e_{i_n} \mid i_j = 1 \text{ or } 2,\ \forall j = 1, 2, \ldots, n\} = H^{\otimes n}. \tag{3.1}$$

In addition, we let Ω be a symbol, and set $\mathbb{C}\Omega = \{\alpha\Omega | \alpha \in \mathbb{C}\}$. Then, $\mathbb{C}\Omega$ is a one-dimensional Hilbert space which is endowed the inner product $\langle \alpha_1\Omega, \alpha_2\Omega \rangle_{\mathbb{C}\Omega} = \alpha_1\overline{\alpha_2}$ where $\alpha_1, \alpha_2 \in \mathbb{C}$. We denote $\mathbb{C}\Omega$ as $H^{\otimes 0}$. Now, we let

$$\mathcal{F}(H) = \bigoplus_{n=0}^{\infty} H^{\otimes n} = \{(h_0, h_1, h_2, \ldots) | h_n \in H^{\otimes n}\ (0 \leq n < \infty) \text{ and } \sum_{n=0}^{\infty} \|h_n\|^2_{H^{\otimes n}} < \infty\}.$$

First, we show that $\mathcal{F}(H)$ is a vector space over $\mathbb{C}$ by checking that $\mathcal{F}(H)$ is closed under vector addition and scalar multiplication (other conditions are obvious.) Given two elements $h^{(1)} = (h_0^{(1)}, h_1^{(1)}, \ldots)$, $h^{(2)} = (h_0^{(2)}, h_1^{(2)}, \ldots) \in \mathcal{F}(H)$, and $\lambda \in \mathbb{C}$. Then, $\|h_j^{(1)} + h_j^{(2)}\|^2_{H^{\otimes j}} \leq (\|h_j^{(1)}\|_{H^{\otimes j}} + \|h_j^{(2)}\|_{H^{\otimes j}})^2 \leq 2(\|h_j^{(1)}\|^2_{H^{\otimes j}} + \|h_j^{(2)}\|^2_{H^{\otimes j}})$ for $0 \leq j < \infty$, and Hence,

$$\sum_{j=0}^{\infty} \|h_j^{(1)} + h_j^{(2)}\|^2_{H^{\otimes j}} < \sum_{j=0}^{\infty} 2(\|h_j^{(1)}\|^2_{H^{\otimes j}} + \|h_j^{(2)}\|^2_{H^{\otimes j}}) < \infty,$$

which implies that $h^{(1)} + h^{(2)} \in \mathcal{F}(H)$.

Since $\|\lambda h_j^{(1)}\|^2_{H^{\otimes j}} = |\lambda|^2 \|h_j^{(1)}\|^2_{H^{\otimes j}}$ for $0 \leq j < \infty$, we have

$$\sum_{j=0}^{\infty} \|\lambda h_j^{(1)}\|^2_{H^{\otimes j}} = |\lambda|^2 \sum_{j=0}^{\infty} \|h_j^{(1)}\|^2_{H^{\otimes j}} < \infty.$$

Thus, $\lambda h^{(1)} \in \mathcal{F}(H)$.

Moreover, consider the map $\langle\ ,\ \rangle : \mathcal{F}(H) \times \mathcal{F}(H) \to \mathbb{C}$ which is defined by

$$\langle (h_0, h_1, \ldots)\ , (g_0, g_1, \ldots) \rangle = \sum_{j=0}^{\infty} \langle h_j, g_j \rangle_{H^{\otimes j}},$$

where $h_j, g_j \in H^{\otimes j}$ for all $0 \leq j < \infty$. It is easy to see that $\langle\ ,\ \rangle$ is an inner product on $\mathcal{F}(H)$, and hence $\mathcal{F}(H)$ is an inner product space.

Finally, we show that $\mathcal{F}(H)$ is a Hilbert space by proving that the norm induced by the inner product is complete. Given $\varepsilon > 0$. Suppose that $\{h^{(n)}\} \subset \mathcal{F}(H)$ is Cauchy. Then there is $N \in \mathbb{N}$ such that for all $n, m \geq N$, $\|h^{(n)} - h^{(m)}\| < \varepsilon$. That is,

$$\sum_{j=0}^{\infty} \|h_j^{(n)} - h_j^{(m)}\|^2_{H^{\otimes j}} < \varepsilon^2.$$

Thus, $\|h_j^{(n)} - h_j^{(m)}\|^2_{H^{\otimes j}} < \varepsilon^2 \Rightarrow \|h_j^{(n)} - h_j^{(m)}\|_{H^{\otimes j}} < \varepsilon$. It follows that for $0 \leq j < \infty$, $\{h_j^{(n)}\}$ is Cauchy in $H^{\otimes j}$. Since $H^{\otimes j}$ is complete, there is $a_j \in H^{\otimes j}$ such that $\lim_{k \to \infty} h_j^{(n)} = a_j$.

Now, let $a = (a_1, a_1, \ldots)$, we claim that $\lim_{n \to \infty} h^{(n)} = a$. Given $\varepsilon > 0$, there is $N \in \mathbb{N}$ such that for $n, m \geq N$, $\|h^{(n)} - h^{(m)}\|^2 < \varepsilon^2/2$.

Fixed $n \geq N$, note that $\|h_j^{(n)} - a_j\|^2_{H^{\otimes j}} = \lim_{m\to\infty} \|h_j^{(n)} - h_j^{(m)}\|^2_{H^{\otimes j}}$ for each $0 \leq j < \infty$. Thus, for $M \in \mathbb{N}$,

$$\begin{aligned}\sum_{j=0}^{M} \|h_j^{(n)} - a_j\|^2_{H^{\otimes j}} &= \sum_{j=0}^{M} \lim_{m\to\infty} \|h_j^{(n)} - h_j^{(m)}\|^2_{H^{\otimes j}} = \lim_{m\to\infty} \sum_{j=0}^{M} \|h_j^{(n)} - h_j^{(m)}\|^2_{H^{\otimes j}} \\ &\leq \lim_{m\to\infty} \sum_{j=0}^{\infty} \|h_j^{(n)} - h_j^{(m)}\|^2_{H^{\otimes j}} = \lim_{m\to\infty} \|h^{(n)} - h^{(m)}\|^2 \leq \frac{\varepsilon^2}{2}.\end{aligned}$$

Since M is arbitrary chosen, we conclude that $\lim_{n\to\infty} h^{(n)} = a$.

Now, we show that $a \in \mathcal{F}(H)$. Let $n \in \mathbb{N}$ such that

$$\sum_{j=0}^{\infty} \|a_j - h_j^{(n)}\|^2_{H^{\otimes j}} \leq 1.$$

We observe that for $0 \leq j < \infty$, $\|a_j\|^2_{H^{\otimes j}} \leq (\|a_j - h_j^{(n)}\|_{H^{\otimes j}} + \|h_j^{(n)}\|_{H^{\otimes j}})^2 \leq 2(\|a_j - h_j^{(n)}\|^2_{H^{\otimes j}} + \|h_j^{(n)}\|^2_{H^{\otimes j}})$. Then,

$$\|a\|^2 = \sum_{j=0}^{\infty} \|a_j\|^2_{H^{\otimes j}} \leq \sum_{j=0}^{\infty} 2(\|a_j - h_j^{(n)}\|^2_{H^{\otimes j}} + \|h_j^{(n)}\|^2_{H^{\otimes j}}) \leq 2 + 2\sum_{j=0}^{\infty} \|h_j^{(n)}\|^2_{H^{\otimes j}} < \infty.$$

We call $\mathcal{F}(H)$ the *full Fock space* over H. Let φ be the vector state on $B(\mathcal{F}(H))$ given by the Ω. That is, φ is defined by

$$\varphi(T) = \langle T\Omega, \Omega\rangle, \text{ where } T \in B(\mathcal{F}(H)). \tag{3.2}$$

For $i = 1, 2$, we consider the linear operator l_i on $\mathcal{F}(H)$ which is determined by the formula:

$$\begin{cases} l_i(\Omega) &= e_i \\ l_i(e_{i_1} \otimes e_{i_2} \otimes \ldots \otimes e_{i_n}) &= e_i \otimes e_{i_1} \otimes e_{i_2} \otimes \ldots \otimes e_{i_n} \\ & \quad \text{where } i_j = 1, 2, \text{and } j = 1, 2, \ldots, n. \end{cases}$$

Now, we prove that for $i = 1, 2$, the linear operator l_i is isometric. First, for fixed $i = 1, 2$ and $n \in \mathbb{N}$, consider the element $h_1 \otimes \ldots \otimes h_n$, where $h_j \in H$ for $j = 1, 2, \ldots, n$. Note that each h_j can be written as the form $a_1^{(j)} e_1 + a_2^{(j)} e_2$, where $a_1^{(j)}, a_2^{(j)} \in \mathbb{C}$. We claim that $l_i(h_1 \otimes \ldots \otimes h_n) = e_i \otimes h_1 \otimes \ldots \otimes h_n$ by the following calculation:

$$\begin{aligned} l_i(h_1 \otimes \ldots \otimes h_n) &= l_i[(a_1^{(1)} e_1 + a_2^{(1)} e_2) \otimes \ldots \otimes (a_1^{(n)} e_1 + a_2^{(n)} e_n)] \\ &= l_i(\sum_{j_1,\ldots,j_n = 1 \text{ or } 2} a_{j_1}^{(1)} a_{j_2}^{(2)} \ldots a_{j_n}^{(n)} e_{j_1} \otimes \ldots \otimes e_{j_n}) \\ &= \sum_{j_1,\ldots,j_n = 1 \text{ or } 2} a_{j_1}^{(1)} a_{j_2}^{(2)} \ldots a_{j_n}^{(n)} l_i(e_{j_1} \otimes \ldots \otimes e_{j_n}) \\ &= \sum_{j_1,\ldots,j_n = 1 \text{ or } 2} a_{j_1}^{(1)} a_{j_2}^{(2)} \ldots a_{j_n}^{(n)} e_i \otimes e_{j_1} \otimes \ldots \otimes e_{j_n} \\ &= e_i \otimes (\sum_{j_1,\ldots,j_n = 1 \text{ or } 2} a_{j_1}^{(1)} a_{j_2}^{(2)} \ldots a_{j_n}^{(n)} e_{j_1} \otimes \ldots \otimes e_{j_n}) \\ &= e_i \otimes h_1 \otimes \ldots \otimes h_n. \end{aligned}$$

Hence, by (3.1) we have $l_i(\zeta) = e_i \otimes \zeta$ for each $\zeta \in H^{\otimes n}$. Therefore, $l_i(\xi) = e_i \otimes \xi$ for $\xi \in \mathcal{F}(H)$. Then, note that

$$\|l_i(\xi)\| = \|e_i \otimes \xi\| = \|e_i\|\|\xi\| = \|\xi\|,$$

which implies that l_i is isometric.

We also can verify easily that the adjoint operator of l_i is described by the formula:

$$\begin{cases} l_i^*(\Omega) &= 0 \\ l_i^*(e_{i_1}) &= \langle e_{i_1}, e_i\rangle\Omega \\ l_i^*(e_{i_1} \otimes e_{i_2} \otimes \ldots \otimes e_{i_n}) &= \langle e_{i_1}, e_i\rangle e_{i_2} \otimes \ldots \otimes e_{i_n}, \\ & \quad \text{where } i_j = 1, 2, \text{and } j = 1, 2, \ldots, n. \end{cases}$$

Lemma 31. *$alg(l_1, l_1^*)$ and $alg(l_2, l_2^*)$ are free.*

Proof. We set $\mathcal{A}_i = alg(l_i, l_i^*)$ for $i = 1, 2$. At first, observe that for $i, j = 1, 2$, we have $l_i^* l_i = 1$ and $l_i^* l_j = 0$ if $i \neq j$; therefore,

$$\mathcal{A}_i = span\{l_i^n l_i^{*m} | \text{ where } n, m \text{ are non-negative integers } \}.$$

Then fixed i and let $T \in \mathcal{A}_i$, T can be written as the form:

$$T = \alpha 1 + \sum_{j=1}^{p} \alpha_j l_i^{m(j)} l_i^{*n(j)},$$

where for $p \in \mathbb{N}$, $1 \leq j \leq p$ we have $(m(j), n(j)) \neq (0, 0)$, and $1 = 1_{\mathcal{F}(H)}$.

Note that $\varphi(T) = \alpha$. This is becasue that $\langle l_i^{m(j)} l_i^{*n(j)}\Omega, \Omega\rangle = 0$. As a result, we conclude that

$$\mathcal{A}_i^o = span\{l_i^m l_i^{*n} | (m, n) \neq (0, 0)\} \text{ where } \mathcal{A}_i^o = \{T \in \mathcal{A}_i | \varphi(T) = 0\}.$$

Now, For $n \in \mathbb{N}$, let $i_1, \ldots, i_n \in \{1, 2\}$ such that $i_1 \neq i_2 \neq \ldots \neq i_n$ and some elements $T_j \in \mathcal{A}_{i_j}^o$ where $j = 1, 2, \ldots, n$. Our goal is to show that $\varphi(T_1 T_2 \ldots T_n) = 0$. We can assume without loss of generality that for each $j = 1, \ldots, n$, $T_j = l_{i_j}^{m(j)} l_{i_j}^{*\,n(j)}$ for some $(m(j), n(j)) \neq (0, 0)$.

Case 1. there is $j \in \{1, 2, \ldots, n-1\}$ such that $n(j) \neq 0$ and $m(j+1) \neq 0$. Since $l_{i_j}^* l_{i_{j+1}} = 0$, we have

$$T_j T_{j+1} = (l_{i_j}^{m(j)}) l_{i_j}^{*\,n(j)-1} l_{i_j}^* l_{i_{j+1}} l_{i_{j+1}}{}^{m(j+1)-1} (l_{i_{j+1}}^*)^{n(j+1)} = 0.$$

Therefore, $\varphi(T_1 T_2 \ldots T_n) = 0$.

Case 2. for any $j \in \{1, 2, \ldots, n-1\}$, we have either $n(j) = 0$ or $m(j+1) = 0$. In this case,

$$T_1 \ldots T_n = \prod_{k=1}^{s} (l_{i_k})^{m(k)} \prod_{h=1}^{p} (l_{i_h}^*)^{n(h)} \quad \text{for some } 0 \leq s, p \leq n.$$

Thus, $\varphi(T_1 \ldots T_n) = \langle \prod_{k=1}^{s} (l_{i_k})^{m(k)} \prod_{h=1}^{p} (l_{i_h}^*)^{n(h)}\Omega, \Omega\rangle = 0$. ☐

Now, if we consider complex number z such that $|z| < 1$, and let

$$W_z = (1 - zl_1)^{-1}\Omega = \sum_{n=0}^{\infty} z^n l_1^n \Omega = \Omega + \sum_{n=1}^{\infty} z^n e_1^{\otimes n}.$$

Then

$$l_1 W_z = \sum_{n=0}^{\infty} z^n e_1^{\otimes(n+1)} = \frac{1}{n}(W_z - \Omega), \quad 0 < |z| < 1.$$

Note that

$$l_1^* W_z = \sum_{n=1}^{\infty} z^n e_1^{\otimes(n-1)} = zW_z, \quad |z| < 1.$$

Observe that $(l_1)^{*n} W_z = (l_1^*)^{n-1} zW_z = z(l_1^*)^{n-2} l_1^* W_z = z(l_1^*)^{n-2} zW_z = z^2 (l_1^*)^{n-2} W_z = \ldots z^n W_z$ for $|z| < 1$. Hence, let f be a polynomial in one variable and $a = l_1 + f(l_1^*)$, we have

$$\begin{aligned} aW_z = (l_1 + f(l_1^*))W_z &= \frac{1}{z}(W_z - \Omega) + f(l_1^*)W_z = \frac{1}{z}(W_z - \Omega) + f(z)W_z \\ &= (\frac{1}{z} + f(z))W_z - \frac{1}{z}\Omega, \quad \text{for } 0 < |z| < 1. \end{aligned}$$

Thus,

$$[(\frac{1}{z} + f(z))1 - a]W_z = \frac{1}{z}\Omega \quad \text{for } 0 < |z| < 1.$$

Since $|1/z + f(z)| \longrightarrow \infty$ as $z \longrightarrow 0$, there is $\delta > 0$ with $0 < \delta \leq 1$ such that $|1/z + f(z)| > \|a\|$ whenever $0 < |z| < \delta$. Now, for $0 < |z| < \delta$,

$$\frac{\|a\|}{|\frac{1}{z} + f(z)|} < 1 \Rightarrow 1 - \frac{a}{\frac{1}{z} + f(z)} \text{ is invertible} \Rightarrow (\frac{1}{z} + f(z))1 - a \text{ is invertible.}$$

Hence, $[(1/z + f(z))1 - a]^{-1}\Omega = zW_z$, and therefore we have

$$\begin{aligned} \varphi([(\frac{1}{z} + f(z))1 - a]^{-1}) &= \langle [(\frac{1}{z} + f(z))1 - a]^{-1}\Omega, \Omega\rangle = \langle zW_z, \Omega\rangle = z\langle W_z, \Omega\rangle \\ &= z(\langle \Omega, \Omega\rangle + \langle \sum_{n=1}^{\infty} z^n e_1^{\otimes n}, \Omega\rangle) = z + \lim_{n\to\infty} \sum_{k=1}^{n} \langle z^k e_1^{\otimes k}, \Omega\rangle = z. \end{aligned}$$

If we set $G_a(\lambda) = \varphi((\lambda 1 - a)^{-1})$, then for $|\lambda| > \|a\|$,

$$G_a(\lambda) = \sum_{n=0}^{\infty} \frac{\varphi(a^n)}{\lambda^{n+1}}.$$

Let $\lambda = 1/z + f(z)$, then

$$G_a(\frac{1}{z} + f(z)) = \varphi([(\frac{1}{z} + f(z))1 - a]^{-1}) = z, \quad 0 < |z| < \delta.$$

Therefore, G_a is invertible in a neighborhood of ∞, and

$$G_a^{-1}(z) = \frac{1}{z} + f(z), \quad 0 < |z| < \delta \Rightarrow f(z) = G_a^{-1}(z) - \frac{1}{z}, \quad 0 < |z| < \delta.$$

As a result, we have the following theorem:

Theorem 32. *If f is a polynomial in one variable, and $a = l_1 + f(l_1^*)$, then $R_a = f$.*

Simliarly, $R_b = g$ where $b = l_2 + g(l_2^*)$ and g is a polynomial in one variable. Now, if f, g are polynomials in one variable, we let

$$a = l_1 + f(l_1^*), \quad b = l_2 + g(l_2^*).$$

Since $alg(a,1) \subseteq alg(l_1, l_1^*)$, $alg(b,1) \subseteq alg(l_2, l_2^*)$, Lemma 31 implies that a, b are free. For $z \in \mathbb{C}$, $|z| < \frac{1}{2}$, we have $\|z(l_1 + l_2)\| \leq |z|(\|l_1\| + \|l_2\|) < 1$. Set

$$\rho_z = (1 - z(l_1 + l_2))^{-1}\Omega = \Omega + \sum_{n=1}^{\infty} z^n (l_1 + l_2)^n \Omega = \Omega + \sum_{n=1}^{\infty} (e_1 + e_2)^{\otimes n}.$$

Then,

$$(l_1 + l_2)\rho_z = \sum_{n=0}^{\infty} z^n (e_1 + e_2)^{\otimes(n+1)} = \frac{1}{z}(\rho_z - \Omega), \quad 0 < |z| < \frac{1}{2}.$$

Since $l_1^* l_1 = 1$, $l_1^* l_2 = 0$,

$$\begin{aligned} l_1^* \rho_z &= l_1^* \Omega + l_1^* (\sum_{n=1}^{\infty} z^n (l_1 + l_2)^n \Omega) = \sum_{n=1}^{\infty} z^n l_1^* (l_1 + l_2)^n \Omega \\ &= \sum_{n=1}^{\infty} z^n l_1^* (l_1 + l_2)(l_1 + l_2)^{n-1} \Omega = \sum_{n=1}^{\infty} z^n (l_1 + l_2)^{n-1} \Omega \\ &= z \rho_z. \end{aligned}$$

Similarly, $l_2^* \rho_z = z \rho_z$. Therefore, if $0 < |z| < \frac{1}{2}$,

$$\begin{aligned} (a + b)\rho_z &= (l_1 + f(l_1^*))\rho_z + (l_2 + g(l_2^*))\rho_z \\ &= (l_1 + l_2)\rho_z + f(l_1^*)\rho_z + g(l_2^*)\rho_z \\ &= \frac{1}{z}(\rho_z - \Omega) + f(z)\rho_z + g(z)\rho_z \\ &= (\frac{1}{z} + f(z) + g(z))\rho_z - \frac{1}{z}\Omega. \end{aligned}$$

Hence,

$$[(\frac{1}{z} + f(z) + g(z))1 - (a + b)]\rho_z = \frac{1}{z}\Omega. \;\; 0 < |z| < \frac{1}{2}.$$

Since $|\frac{1}{z} + f(z) + g(z)| \longrightarrow \infty$ as $|z| \longrightarrow \infty$, there is $\delta > 0$ with $0 < \delta \leq 1/2$ such that

$$|\frac{1}{z} + f(z) + g(z)| > \|a + b\|, \quad \text{if } 0 < |z| < \delta.$$

As a result, for $0 < |z| < \delta$,

$$(\frac{1}{z} + f(z) + g(z))1 - (a + b) \text{ is invertible, and } \; [(\frac{1}{z} + f(z) + g(z))1 - (a + b)]^{-1}\Omega = z\rho_z.$$

Hence, for $0 < |z| < \delta$,

$$\begin{aligned} \varphi([(\frac{1}{z} + f(z) + g(z))1 - (a + b)]^{-1}) &= \langle z\rho_z, \Omega \rangle \\ &= z\langle \Omega + \sum_{n=1}^{\infty} z^n (e_1 + e_2)^{\otimes n}, \Omega \rangle \\ &= z(\langle \Omega, \Omega \rangle + \langle \sum_{n=1}^{\infty} z^n (e_1 + e_2)^{\otimes n}, \Omega \rangle) \\ &= z. \end{aligned}$$

By the same arugment as the proof in Theorem 32, we have

$$R_{a+b} = f + g.$$

Moreover, Theorem 32 implies $R_a = f$ and $R_b = g$. As a result, we obtain the following conclusion:

Theorem 33. *If f, g are polynomials in one variable, and $a = l_1 + f(l_1^*)$, $b = l_2 + g(l_2^*)$, then $R_{a+b} = R_a + R_b$.*

Based on Theorem 32 and Theorem 33, we start to prove Theorem 30. That is, to prove that if a, b are free in a non-commutative probability space $(\mathcal{A}, \varphi)$, then $R_{a+b} = R_a + R_b$.

Proof. Firstly, we note that

$$R_a(z) = \sum_{k=0}^{\infty} S_k(a) z^k, \quad R_b(z) = \sum_{k=0}^{\infty} S_k(b) z^k$$

where $\{S_k(a)\}_{k=0}^{\infty}, \{S_k(b)\}_{k=0}^{\infty}$ are sequences in complex numbers. We fix $n \in \mathbb{N}$ and let

$$a' = l_1 + \sum_{k=0}^{n-1} S_k(a) {l_1^*}^k, \quad b' = l_2 + \sum_{k=0}^{n-1} S_k(b) {l_2^*}^k.$$

Then, we write the corresponding R-transform of a' and b' by

$$R_{a'}(z) = \sum_{k=0}^{\infty} S_k(a') z^k, \quad R_{b'}(z) = \sum_{k=0}^{\infty} S_k(b') z^k$$

where $\{S_k(a')\}_{k=0}^{\infty}, \{S_k(b')\}_{k=0}^{\infty}$ are sequences in complex numbers.

Let φ' be the state on $B(\mathcal{F}(H))$, which is defined as (3.2). By Theorem 32, we have

$$R_{a'}(z) = \sum_{k=0}^{n-1} S_k(a) z^k, \quad R_{b'}(z) = \sum_{k=0}^{n-1} S_k(b) z^k,$$

which implies that $S_k(a') = S_k(a)$ and $S_k(b') = S_k(b)$ for $k = 0, 1, \ldots, n-1$. According to the definition of R-transform, we obtain

$$\varphi(a^k) = \varphi'({a'}^k), \text{ and } \varphi(b^k) = \varphi'({b'}^k) \text{ for all } k = 0, 1, \ldots, n.$$

By Corollary 7, for $k = 0, 1, \ldots, n$, we have the following property:

$$\varphi((a+b)^k) = P_{2k}(\varphi(a), \varphi(a^2), \ldots, \varphi(a^k), \varphi(b), \varphi(b^2), \ldots, \varphi(b^k)),$$

where P_{2k} is a polynomial in $2k$ variables. Therefore, we deduce that

$$\varphi((a+b)^k) = \varphi'((a'+b')^k) \text{ for all } k = 0, 1, \ldots n.$$

Hence, $R_{a+b}(z)$ coincides with $R_{a'+b'}(z)$ up to order $n-1$ in z, and therefore, $R_{a+b}(z)$ coincides with $R_a(z) + R_b(z)$ up to order $n-1$ in z. Since n is arbitrary, we get our final result $R_{a+b} = R_a + R_b$. □

In the remaining part of this chapter, we introduce the analyticity of R-transform which is based on the probability measure with compact support on the real line. At first, given probability measure ν which has compact support contains in $[-r, r]$ for some $r > 0$. By Theorem 29, there is a analytic function R_ν on an open disc containing 0 such that

$$G_\nu(R_\nu(z) + \frac{1}{z}) = z, \quad 0 < |z| < \frac{1}{6r}$$

where G_ν is the Cauchy transform of ν. We call R_ν is the R-transform of the probability measure ν, and denote it by R_ν.

Theorem 34. *Let μ, ν be probability measures on $\mathbb{R}$ with compact support and R_μ, R_ν be the corresponding R-transforms. Then there is a unique compactly supported probability measure m on $\mathbb{R}$ such that $R_m = R_\mu + R_\nu$, where R_m is the R-transform of m.*

Proof. Let μ, ν be two probability measures with compact support on $\mathbb{R}$. we set

$$\mathcal{A}_1 = \{f | f : supp(\mu) \to \mathbb{C} \text{ is continuous}\} \text{ and } \mathcal{A}_2 = \{f | f : supp(\nu) \to \mathbb{C} \text{ is continuous}\}.$$

It is easy to see that $\mathcal{A}_1, \mathcal{A}_2$ are unital C^*-algebras. For $i = 1, 2$, we define $\varphi_i : \mathcal{A}_i \to \mathbb{C}$ by $\varphi_1(f) = \int_{\mathbb{R}} f(t)d\mu(t)$, and $\varphi_2(f) = \int_{\mathbb{R}} f(t)d\nu(t)$ Then, $(\mathcal{A}_i, \varphi_i)$ becomes a C^*-probability space and φ_i is obviously faithful and trace for $i = 1, 2$.

Now, we consider $a \in \mathcal{A}_1$ be the identity function on $supp(\mu)$, and $b \in \mathcal{A}_2$ be the identity function on $supp(\nu)$. Then it is obvious that

$$\varphi_1(a^n) = \int_{\mathbb{R}} t^n d\mu(t), \ \varphi_2(b^n) = \int_{\mathbb{R}} t^n d\nu(t), \ n \geq 0.$$

We let $(\mathcal{A}, \varphi)$ be the free product of C^*-probability spaces $(\mathcal{A}_1, \varphi_1)$ and $(\mathcal{A}_2, \varphi_2)$. Then $(\mathcal{A}, \varphi)$ is a C^*-probability space, and there are norm preserving unital $*$-homomorphisms $W_i : \mathcal{A}_i \to \mathcal{A}$ for $i = 1, 2$ such that $W_1(\mathcal{A}_1)$ and $W_2(\mathcal{A}_2)$ are free in $(\mathcal{A}, \varphi)$. We set $W_1(a) = \tilde{a}$, and $W_2(b) = \tilde{b}$. We observe that $\tilde{a} + \tilde{b} \in \mathcal{A}$ is self-adjoint; in addition,

$$\varphi(\tilde{a}^n) = \varphi_1(a^n) = \int_{\mathbb{R}} t^n d\mu(t), \ \varphi(\tilde{b}^n) = \varphi_2(b^n) = \int_{\mathbb{R}} t^n d\nu(t), \ n \geq 0.$$

On the other hand, since $\tilde{a} + \tilde{b}$ is self-adjoint, there is a unique probability measure m with compact support on $\mathbb{R}$ such that

$$\varphi((\tilde{a} + \tilde{b})^n) = \int_{\mathbb{R}} t^n dm(t), \ n \geq 0.$$

Finally, by Theorem 30, we have $R_{\tilde{a}} + R_{\tilde{b}} = R_{\tilde{a}+\tilde{b}} = R_{\mu \boxplus \nu}$. However, $R_\mu = R_a = R_{\tilde{a}}$ and $R_\nu = R_b = R_{\tilde{b}}$. Hence, we complete the proof. □

Definition 35. The unique probability measure m as in Theorem 34 is denoted by $\mu \boxplus \nu$, and we call it the *free convolution* of μ and ν.

We now calculate the following example. At first, we recall that a real symmetric $n \times n$ random matrix $X_n = [a_{ij}]$ is called a Wigner matrix if (i) the family $\{a_{ii} : i \in \mathbb{N}\}$ of diagonal entries is independent and identically distributed (i.i.d.) and so is the family $\{a_{ij} : i, j \in \mathbb{N}, i < j\}$ of variables in the upper diagonal part of our matrix ensemble, (ii) the diagonal family $\{a_{ii} : i \in \mathbb{N}\}$ is independent from the upper diagonal family $\{a_{ij} : i, j \in \mathbb{N}, i < j\}$, (iii) $Ea_{ij} = 0$ and $Ea_{ij}^2 = 1$.

Example 36. *Let X_n and Y_n be two $n \times n$ independent (all entries of X_n are independent to that of Y_n) Wigner matrices of Gaussian entries. Note that Y_n is invariant to orthogonal conjugation; that is, for any $n \times n$ orthogonal matrix O, the joint probability distribution of the entries in the orthogonal conjugation OY_nO^* is the same as the joint probability distribution of the original entries in Y_n. Recall from Voiculescu's asymptotic freeness result [6], $\frac{1}{\sqrt{n}}X_n + \frac{1}{\sqrt{n}}Y_n$ converges weakly in probability to the free convolution $\gamma \boxplus \gamma$, where γ is the semicircle law defined by*

$$d\gamma(t) = \frac{1}{2\pi}\sqrt{4-t^2}dt, \quad t \in [-2, 2].$$

We now use Theorem 34 to compute the measure $\gamma \boxplus \gamma$. At first, we write

$$C_p = \frac{1}{p+1}\binom{2p}{p}, \ p \geq 0,$$

for the p-th Catalan number. Note that Catalan numbers have the recurrence relation

$$\begin{cases} C_0 &= C_1 = 1 \\ C_p &= \sum_{j=1}^{p} C_{j-1}C_{p-j}, \ p \geq 2. \end{cases}$$

Let G_γ be the Cauchy transform of γ. Observe that

$$\int_{\mathbb{R}} t^k d\gamma(t) = \begin{cases} 0 & \text{if } k \text{ is odd} \\ C_p & \text{if } k \text{ is even, } k = 2p, \end{cases} \tag{3.3}$$

For $|z|$ sufficiently large, we have

$$G_\gamma(z) = \sum_{p=0}^{\infty} \frac{C_p}{z^{2p+1}}.$$

By (3.3) and some elementary manipulations of power series, we have

$$\begin{aligned} G_\gamma(z) &= \frac{1}{z} + \sum_{p=1}^{\infty} \frac{1}{z^{2p+1}}(\sum_{j=1}^{p} C_{j-1}C_{p-j}) = \frac{1}{z} + \frac{1}{z}\sum_{p=1}^{\infty}\sum_{j=1}^{p} \frac{C_{j-1}}{z^{2j-1}}\frac{C_{p-j}}{z^{2(p-j)+1}} \\ &= \frac{1}{z}\sum_{j=1}^{\infty}\frac{C_{j-1}}{z^{2j-1}}(\sum_{p=j}^{\infty}\frac{C_{p-j}}{z^{2(p-j)+1}}) = \frac{1}{z} + \frac{1}{z}\sum_{j=1}^{\infty}\frac{C_{j-1}}{z^{2j-1}}G_\gamma(z) \\ &= \frac{1}{z} + \frac{1}{z}G_\gamma(z)^2. \end{aligned}$$

Thus, G_γ satisfies that equation $G_\gamma(z)^2 - zG_\gamma(z) + 1 = 0, \ z \in \mathbb{C}^+$. It is easy to solve that

$$G_\gamma(z) = \frac{z \pm \sqrt{z^2-4}}{2}.$$

We need to choose a branch of $\sqrt{z^2-4}$ *for* $z \in \mathbb{C}^+$. *We write* $z^2 - 4 = (z-2)(z+2)$ *and define each of* $\sqrt{z-2}$ *and* $\sqrt{z+2}$ *on* $\mathbb{C}^+$. *For* $z \in \mathbb{C}^+$, *let* θ_1 *be the angle between the x-axis and the line joining* z *to* 2*; and* θ_2 *the angle between the x-axis and the line joining* z *to* -2. *Then,* $z - 2 = |z-2|e^{i\theta_1}$ *and* $z+2 = |z+2|e^{i\theta_2}$, *so* $\sqrt{z^2-4}$ *is defined by* $|z^2-4|^{1/2}e^{i(\theta_1+\theta_2)/2}$.

Note that

$$\lim_{y\to\infty}(iy)\frac{iy-\sqrt{(iy)^2-4}}{2} = 1 \text{ and } \lim_{y\to\infty}(iy)\frac{iy+\sqrt{(iy)^2-4}}{2} = \infty.$$

Thus, $\lim_{y\to\infty} iyG_\gamma(iy) = 1$ *by Proposition 26 implies that*

$$G_\gamma(z) = \frac{z-\sqrt{z^2-4}}{2}.$$

Note that $G_\gamma^{-1}(z)$ *exists for* $|z|$ *sufficiently large. Let* $w = G_\gamma(z)$, *then*

$$w = \frac{z-\sqrt{z^2-4}}{2} \Rightarrow (2w-z)^2 = (\sqrt{z^2-4})^2 \Rightarrow w^2 - zw + 1 = 0 \Rightarrow z = w + \frac{1}{w} = G_\gamma^{-1}(w).$$

Thus, the R-transform of γ *is* $R_\gamma(w) = w$. *We obtain* $R_{\gamma\boxplus\gamma}(w) = R_\gamma(w) + R_\gamma(w) = 2w$ *by Theorem 34. Since* $R_{\gamma\boxplus\gamma}(w) = G^{-1}_{\gamma\boxplus\gamma}(w) - 1/w$, *we have* $G^{-1}_{\gamma\boxplus\gamma}(w) = 2w + 1/w$. *Hence, the Cauchy transform of* $\gamma \boxplus \gamma$ *is*

$$G_{\gamma\boxplus\gamma}(z) = \frac{z-\sqrt{z^2-8}}{4}.$$

Now, we review the Stieltjes inversion formula (Theorem 28): if G_ν *is the Cauchy transform of some probability measure* ν, *then for* $x \in \mathbb{R}$

$$d\nu(x) = \lim_{y\downarrow 0^+} \frac{-1}{\pi} Im(G_\nu(x+iy))dx.$$

The latter limit is considered in the weak topology on the space of probability measure on $\mathbb{R}$.

Thus for $x \in (-\sqrt{8}, \sqrt{8})$ *and* $y > 0$, *we have*

$$\frac{-1}{\pi} Im(G_{\gamma\boxplus\gamma}(x+iy))dx = \frac{-1}{\pi} Im(\frac{x+iy+\sqrt{(x+iy)^2-8}}{4})dx \to \frac{1}{4\pi}\sqrt{8-x^2} \quad as \quad y \downarrow 0^+.$$

As a result, we get

$$d\gamma \boxplus \gamma(x) = \frac{1}{4\pi}\sqrt{8-x^2}dx, \quad x \in (-\sqrt{8}, \sqrt{8}).$$

Observe that the first moment of $d\gamma \boxplus \gamma$ *is zero, and*

$$\int_{\mathbb{R}} x^2 d\gamma \boxplus \gamma(x) = \int_{-\sqrt{8}}^{\sqrt{8}} x^2 \frac{1}{4\pi}\sqrt{8-x^2}dx = \left.\frac{\sqrt{8-x^2}(x^3-4x) + 32\sin^{-1}(x/2\sqrt{2})}{4}\right|_{-\sqrt{8}}^{\sqrt{8}} = 2.$$

Therefore, the free convolution $\gamma \boxplus \gamma$ *is another semicircle with variance* 2.

Chapter 4

Free Convolution: Measures with Finite Variance

In Chapter 3, we proved that if ν is a probability measure on the real line with compact support, then the R-transform of ν is analytic on an open disc containing 0. Furthermore, if we have two compactly supported probability measures μ, ν on $\mathbb{R}$, then there exists the compactly supported probability measure $\mu \boxplus \nu$ on $\mathbb{R}$ satisfies the R-transform addition formula.

In this chapter, we will show that if ν is a measure with finite variance on the real line, but make no assumption about the support, then we still have the similar result: there is an analytic function R_ν satisfies the equation $G_\nu(R_\nu(z) + 1/z) = z$ on an open disc in the lower half plane. Moreover, if μ, ν are probability measures on $\mathbb{R}$, then there is a probability measure $\mu \boxplus \nu$ such that

$$R_{\mu\boxplus\nu} = R_\mu + R_\nu,$$

where $R_{\mu\boxplus\nu}$, R_μ, and R_ν are the R-transforms of $\mu \boxplus \nu$, μ, and ν.

Lemma 37. *Let ν be a probability measure on $\mathbb{R}$ and G_ν its Cauchy transform. Let $F_\nu(z) = 1/G_\nu(z)$. Then F_ν maps $\mathbb{C}^+$ to $\mathbb{C}^+$ and $Im(z) \le Im(F_\nu(z))$ for all $z \in \mathbb{C}^+$, with equality for some z only if ν is a Dirac mass.*

Proof. For $z \in \mathbb{C}^+$,

$$G_\nu(z) = \int_{\mathbb{R}} \frac{\overline{z-t}}{z-t}\frac{1}{\overline{z-t}}d\nu(t) = \int_{\mathbb{R}} \frac{\bar{z}-t}{|z-t|^2}d\nu(t).$$

Then $Im(G_\nu(z)) = -Im(z)\int_{\mathbb{R}} 1/|z-t|^2 d\nu(t)$ and $Im(F_\nu(z)) = Im(\overline{G_\nu(z)}/|G_\nu(z)|^2)$. Therefore, we have

$$\frac{Im(F_\nu(z))}{Im(z)} = -\frac{1}{Im(z)}\frac{1}{|G_\nu(z)|^2}Im(G_\nu(z)) = \frac{\int_{\mathbb{R}} 1/|z-t|^2 d\nu(t)}{|G_\nu(z)|^2}. \tag{4.1}$$

Now, it suffices to show that the right hand side of (4.1) is less than or equal to 1. By Cauchy Schwarz inequality,

$$|G_\nu(z)|^2 = |\int_{\mathbb{R}} \frac{1}{z-t}d\nu(t)|^2 \le \int_{\mathbb{R}} 1 d\nu(t) \int_{\mathbb{R}} |\frac{1}{z-t}|^2 d\nu(t) = \int_{\mathbb{R}} |\frac{1}{z-t}|^2 d\nu(t).$$

Thus, (4.1) $\ge$ 1.

(4.1) $= 1$ only if $t \to 1/(z-t)$ is $\nu-$ almost everywhere. That is, ν is a Dirac mass. $\square$

Lemma 38. *Let ν be a probability measure with finite variance σ^2 and let $G_1(z) = z - 1/G_\nu(z)$, where G_ν is the Cauchy transform of ν. Then there is a probability measure ν_1 such that*

$$G_1(z) = \alpha_1 + \sigma^2 \int_{\mathbb{R}} \frac{1}{z-t}d\nu_1(t)$$

where α_1 is the mean of ν.

Proof. If $\sigma^2 = 0$, then $\int t^2 d\nu(t) = (\int t d\nu(t))^2$, which impies that ν is a Dirac mass by Jensen's inequality. Thus, the result is true obviously. Assume that $\sigma^2 \neq 0$. Note that $1/G_\nu$ is analytic on $\mathbb{C}^+$ since the analytic function G_ν is nowhere zero on $\mathbb{C}^+$, and so is $G_1(z) = z - 1/G_\nu(z)$. By Lemma 37, we have $G_1(\mathbb{C}^+) \subset \mathbb{C}^-$. Also, $(G_1 - \alpha_1)(\mathbb{C}^+) \subset \mathbb{C}^-$. Now, let α_1, α_2 be the first and second, respectively, moment of ν. We note that

$$G_1(z) = z - \frac{1}{G_\nu(z)} = z - \frac{1}{1/z + \alpha_1/z^2 + \alpha_2/z^3 + o(1/z^3)} = \alpha_1 + \frac{\alpha_2 - \alpha_1^2}{z} + o(1/z). \tag{4.2}$$

As a result, $\lim_{z\to\infty} z(G_1(z) - \alpha_1) = \alpha_2 - \alpha_1^2 = \sigma^2 > 0$. By Theorem 27, there is a probability measure ν_1 such that

$$G_1(z) - \alpha_1 = \sigma^2 \int_{\mathbb{R}} \frac{1}{z-t} d\nu_1(t).$$

$\square$

Notation 39. *Let $\alpha > 0$ and let $\Gamma_\alpha = \{x + yi | \alpha y > |x|\}$ and for $\beta > 0$ let $\Gamma_{\alpha,\beta} = \{z \in \Gamma_\alpha | Im(z) > \beta\}$. We call Γ_α a Stolz angle, and $\Gamma_{\alpha,\beta}$ a truncated Stolz angle. Note that for $z \in \mathbb{C}^+$, we have $z \in \Gamma_\alpha$ if and only if $\sqrt{1+\alpha^2} Im(z) > |z|$.*

If $\alpha > 0$ and f is a function on Γ_α, we say that $\lim_{z\to\infty} f(z) = c$ in Stolz angle α if $\forall \varepsilon > 0$, $\exists\, \beta > 0$ such that $|f(z) - c| < \varepsilon$ for $z \in \Gamma_{\alpha,\beta}$. We denote it by $\lim_{z\to\infty, z\in\Gamma_\alpha} f(z) = c$.

Lemma 40. *Let ν be a probability measure on $\mathbb{R}$ and $\alpha > 0$. Then*
(1) For $z \in \Gamma_\alpha$ and $t \in \mathbb{R}$, $|z - t| \geq |t|/\sqrt{1+\alpha^2}$,
(2) For $z \in \Gamma_\alpha$ and $t \in \mathbb{R}$, $|z - t| \geq |z|/\sqrt{1+\alpha^2}$,
(3) $\lim_{z\to\infty, z\in\Gamma_\alpha} \int_{\mathbb{R}} t/(z-t) d\nu(t) = 0$,
(4) $\lim_{z\to\infty, z\in\Gamma_\alpha} zG_\nu(z) = 1$.

Proof. (1) The case $t = 0$ is obvious, and if $t > 0$, then the distance between $(t, 0)$ and $z \in \Gamma_\alpha$ is

$$\sqrt{(t - \frac{t\alpha^2}{\alpha^2+1})^2 + (0 - \frac{t\alpha}{\alpha^2+1})^2} = \sqrt{\frac{((\alpha^2+1)t - t\alpha^2)^2 + t^2\alpha^2}{(\alpha^2+1)^2}} = \frac{t}{\sqrt{1+\alpha^2}}.$$

Hence, $|z - t| \geq t/\sqrt{1+\alpha^2}$ if $t > 0$, and thus, $|z - t| \geq |t|/\sqrt{1+\alpha^2}$ for all $t \in \mathbb{R}$.

(2) For $z \in \Gamma_\alpha$, we write $z = |z|e^{i\theta}$ with $tan^{-1}(\alpha^{-1}) < \theta < \pi - tan^{-1}(\alpha^{-1})$. If $t = 0$, it is trivial. If $t > 0$, then

$$|z - t| = |\overline{z} - t| = ||z|e^{-i\theta} - t| = ||z| - te^{i\theta}| \geq \frac{|z|}{\sqrt{1+\alpha^2}} \quad \text{since } te^{i\theta} \in \Gamma_\alpha.$$

For $t < 0$,

$$|z - t| = ||z|e^{i\theta} - t| = ||z| - te^{-i\theta}| \geq \frac{|z|}{\sqrt{1+\alpha^2}} \quad \text{since } te^{-i\theta} \in \Gamma_\alpha.$$

(3) Since $|t/(z-t)| \leq \sqrt{1+\alpha^2}$ for all $z \in \Gamma_\alpha$, by Dominated Convergence Theorem we have

$$\lim_{z\to\infty; z\in\Gamma_\alpha} \int_{\mathbb{R}} \frac{t}{z-t} d\nu(t) = \int_{\mathbb{R}} 0 d\nu(t) = 0.$$

(4) Note that $zG_\nu(z) - 1 = \int_{\mathbb{R}} t/(z-t) d\nu(t)$, then we have $\lim_{z\to\infty; z\in\Gamma_\alpha} (zG_\nu(z) - 1) = 0$. $\square$

Lemma 41. *Suppose $F : \mathbb{C}^+ \to \mathbb{C}^+$ is analytic and there is $C > 0$ such that $|F(z) - z| \le C/Im(z)$ for all $z \in \mathbb{C}^+$. Then there is a probability measure ν with mean 0 and variance $\sigma^2 \le C$ such that $1/F$ is the Cauchy transform of ν. Furthermore, σ^2 is the smallest C such that $|F(z) - z| \le C/Im(z)$.*

Proof. Let $G(z) = 1/F(z)$, then $G : \mathbb{C}^+ \to \mathbb{C}^-$ is analytic and

$$|1 - \frac{1}{zG(z)}| = |\frac{z}{z} - \frac{F(z)}{z}| = \frac{|z - F(z)|}{|z|} \le \frac{C}{|z|Im(z)}.$$

Thus, for $\alpha, \beta > 0$, we consider $z \in \Gamma_{\alpha,\beta}$, then

$$\begin{aligned} |zG(z) - 1| \le \frac{C}{Im(z)}|G(z)| &\le \frac{C}{Im(z)}(\int_{\mathbb{R}} \frac{1}{|z-t|^2} d\nu(t))^{1/2} \\ &\le \frac{C}{Im(z)}(\int_{\mathbb{R}} \frac{1}{Im(z)^2} d\nu(t))^{1/2} = \frac{C}{Im(z)}\frac{1}{Im(z)} = \frac{C}{(Im(z))^2}. \end{aligned}$$

As a result, we have $\lim_{z\to\infty} zG(z) = 1$ in any Stolz angle. In particular,

$$|\ y|G(iy)| - 1| \le |\ |iy||G(iy)| - 1| \le |(iy)G(iy) - 1| \to 0 \text{ as } y \to \infty.$$

Thus, by Theorem 27, there is a probability measure ν such that G is the Cauchy transform of ν. Now,

$$\begin{aligned} yIm[iyG(iy)(F(iy) - iy)] &= y^2 + y^2yIm(G(iy)) = y^2 + (-y^4)\int_{\mathbb{R}} \frac{d\nu(t)}{y^2+t^2} \\ &= y^2[1 - \int_{\mathbb{R}} \frac{y^2}{y^2+t^2} d\nu(t)] = \int_{\mathbb{R}} \frac{y^2}{y^2+t^2} t^2 d\nu(t). \end{aligned}$$

Since $(y^2t^2)/(y^2+t^2)$ is increasing that converges to t^2 as $y \to \infty$, by Monotone Convergence Theorem,

$$\int_{\mathbb{R}} t^2 d\nu(t) = \lim_{y\to\infty} \int_{\mathbb{R}} \frac{y^2}{y^2+t^2} t^2 d\nu(t).$$

However,

$$\begin{aligned} |yIm[iyG(iy)(F(iy) - iy)]| &= |yIm(iyG(iy))||Im(F(iy) - iy)| \\ &\le y|iyG(iy)||F(iy) - iy| \\ &\le y|iyG(iy)|\frac{C}{Im(iy)} \\ &\le \frac{y|iyG(iy)|C}{Im(iy)} = C|iyG(iy)| \end{aligned}$$

and $\lim_{y\to\infty} |iyG(iy)| = 1$, we conclude that $\int_{\mathbb{R}} t^2 d\nu(t) \le C$. But, $\sigma^2 \le C$, which implies that ν has the first and second moments. Now, we note that

$$-Re[iyG(iy)(F(iy) - iy)] = -y^2Re(G(iy)) = \int_{\mathbb{R}} \frac{y^2}{y^2+t^2} t d\nu(t).$$

Since $\lim_{y\to\infty} iyG(iy) = 1$ and $|F(iy) - iy| \le C/y$,

$$|Re[iyG(iy)(F(iy) - iy)]| \le |iyG(iy)||F(iy) - iy| \to 0 \text{ as } y \to \infty.$$

Hence, by Monotone Convergence Theorem, the first moment of ν is 0.

We now have that $\sigma^2 \leq C$. We will claim that σ^2 is the smallest C such that $|F(z) - z| \leq C/Im(z)$. At first, we consider ν is a Dirac mass, then $\nu = \delta_0$ since the mean of ν is zero. Thus, $F(z) = z$, which implies that the minimal C is $0 = \sigma^2$.

We assume that ν is not a Dirac mass. For $z \in \mathbb{C}^+$, we have $z - F(z) \in \mathbb{C}^-$ by Lemma 37. On the other hand, by equation (4.2),

$$\lim_{z\to\infty}(z - F(z)) = \lim_{z\to\infty}(\sigma^2 + o(\frac{1}{z})) = \sigma^2 \text{ in any Stolz angle.}$$

By Theorem 27, there is a probability measure $\widetilde{\nu}$ such that

$$z - F(z) = \sigma^2 \int_{\mathbb{R}} \frac{1}{z-t} d\widetilde{\nu}(t).$$

Hence, we have

$$|z - F(z)| \leq \sigma^2 \int_{\mathbb{R}} \frac{1}{|z-t|} d\widetilde{\nu}(t) \leq \int_{\mathbb{R}} \frac{\sigma^2}{Im(z)} d\widetilde{\nu}(t) = \frac{\sigma^2}{Im(z)}.$$

This prove the last claim. □

Notation 42. *For $\beta > 0$, let $\mathbb{C}^+_\beta = \{z | Im(z) > \beta\}$.*

Lemma 43. *Suppose $F : \mathbb{C}^+ \to \mathbb{C}^+$ is analytic and there is $\sigma > 0$ such that $|z - F(z)| \leq \sigma^2/Im(z)$ for $z \in \mathbb{C}^+$. Then*
(1) $\mathbb{C}^+_{2\sigma} \subset F(\mathbb{C}^+_\sigma)$,
(2) for each $w \in \mathbb{C}^+_{2\sigma}$, there is a unique $z \in \mathbb{C}^+_\sigma$ such that $F(z) = w$.
Hence, there is an analytic function F^{-1}, defined on $\mathbb{C}^+_{2\sigma}$ such that $F(F^{-1}(w)) = w$. Moreover, for $w \in \mathbb{C}^+_{2\sigma}$
(3) $Im(F^{-1}(w)) \leq Im(w) \leq 2Im(F^{-1}(w))$, and
(4) $|F^{-1}(w) - w| \leq 2\sigma^2/Im(w)$.

Proof. (1) Given $w \in \mathbb{C}^+_{2\sigma}$, we let $z \in \mathbb{C}$ such that $|z - w| = \sigma$. Then,

$$Im(z) \geq Im(w - i\sigma) = Im(w) - \sigma > 2\sigma - \sigma = \sigma.$$

Let C be the circle with centre w and radius σ. Then, $C \subseteq \mathbb{C}^+_{2\sigma}$. For $z \in C$, we have

$$|(F(z) - w) - (z - w)| = |F(z) - z| = \frac{\sigma^2}{Im(z)} < \sigma = |z - w|,$$

and note that $F(z) - w,\ z - w$ have the same number of roots inside C. By Rouché's Theorem, there is a unique $z \in int(C)$ such that $F(z) = w$. Thus, $w \in F(\mathbb{C}^+_\sigma)$.
(2) Let $z' \in \mathbb{C}^+_\sigma$ with $F(z') = w$, then

$$|w - z'| = |F(z') - z'| \leq \frac{\sigma^2}{Im(z')} < \sigma.$$

So $z' \in Int(C)$, and hence $z = z'$. This proves (2).

Since F is analytic and injective on the inverse image of $\mathbb{C}^+_{2\sigma}$, F has non-zero derivative on this region. In particular, the inverse of F defined on $\mathbb{C}^+_{2\sigma}$ is analytic.

(3) By Lemma 41, $1/F$ is the Cauchy transform of a probability measure with finite variance. Then, by Lemma 37, we have $Im(F(z)) \geq Im(z)$, which implies that $Im(w) \geq Im(F^{-1}(w))$.

Now, we note that if we replace σ in (1) by $\beta > \sigma$, then for $w \in \mathbb{C}^+_{2\beta}$, $Im(F^{-1}(w)) > \beta$. Let $2\beta \nearrow Im(w)$, we have $Im(F^{-1}(w)) \geq 1/2Im(w)$.

(4) For $w \in \mathbb{C}^+_{2\sigma}$, let $z = F^{-1}(w) \in \mathbb{C}^+_\sigma$. By (3), we have $Im(z) \geq 1/2Im(w)$, which implies that $2Im(z) \geq Im(w)$. Then

$$|F^{-1}(w) - w| = |z - w| = |F(z) - z| \leq \frac{\sigma^2}{Im(z)} \leq \frac{2\sigma^2}{Im(w)}.$$

□

Now, we consider ν be a probability measure on $\mathbb{R}$ with first and second moments α_1, α_2, and let G_ν be the corresponding Cauchy transform of ν and σ^2 be the variance of ν.

Theorem 44. *Let $F_\nu = 1/G_\nu$, then we have $|F_\nu(z) + \alpha_1 - z| \leq \sigma^2/Im(z)$. Furthermore, there is an analytic function G_ν^{-1} defined on $\{z||z + i(4\sigma)^{-1}| < 1/(4\sigma)\}$ such that $G_\nu(G_\nu^{-1}(z)) = z$.*

Proof. By Lemma 38, there is a probability measure ν_1 such that

$$z - F_\nu(z) = \alpha_1 + \sigma^2 \int_{\mathbb{R}} \frac{1}{z-t} d\nu_1(t).$$

Hence,

$$|F_\nu(z) + \alpha_1 - z| \leq \int_{\mathbb{R}} \frac{\sigma^2}{|z-t|} d\nu_1(t) \leq \int_{\mathbb{R}} \frac{\sigma^2}{Im(z)} d\nu_1(t) = \frac{\sigma^2}{Im(z)}.$$

Thus, $|F_\nu(z) + \alpha_1 - z| \leq \sigma^2/Im(z)$.

Let $\widetilde{G_\nu}(z) = G_\nu(z + \alpha_1)$. Then

$$\widetilde{G_\nu}(z) = \int_{\mathbb{R}} \frac{1}{z - (t - \alpha_1)} d\nu(t)$$

is the Cauchy transform of a centered probability measure. Let $\widetilde{F_\nu}(z) = 1/\widetilde{G_\nu}$ then $\widetilde{F_\nu} : \mathbb{C}^+ \to \mathbb{C}^+$. By Lemma 38, there is a probability measure $\widetilde{\nu}$ such that $z - \widetilde{F_\nu}(z) = \sigma^2 \int (z-t)^{-1} d\widetilde{\nu}(t)$. Hence,

$$|z - \widetilde{F_\nu}(z)| \leq \int_{\mathbb{R}} \frac{\sigma^2}{|z-t|} d\widetilde{\nu}(t) \leq \int_{\mathbb{R}} \frac{\sigma^2}{Im(z)} d\widetilde{\nu}(t) = \frac{\sigma^2}{Im(z)}.$$

By Lemma 43, we get an inverse function $\widetilde{F_\nu}^{-1}$ for $\widetilde{F_\nu}$ on $\{z|Im(z) > 2\sigma\}$. Note that $|z + i(4\sigma)^{-1}| < 1/(4\sigma)^{-1}$ if and only if $Im(1/z) > 2\sigma$. Since $G_\nu(z) = \widetilde{G_\nu}(z - \alpha_1) = 1/\widetilde{F_\nu}(z - \alpha_1)$, we let $G_\nu^{-1}(z) = \widetilde{F_\nu}^{-1}(1/z) + \alpha_1$ for $|z + i(4\sigma)^{-1}| < (4\sigma)^{-1}$. Then,

$$G_\nu(G_\nu^{-1}(z)) = G_\nu(\widetilde{F_\nu}^{-1}(1/z) + \alpha_1) = \widetilde{G_\nu}(\widetilde{F_\nu}^{-1}(1/z)) = \frac{1}{\widetilde{F_\nu}(\widetilde{F_\nu}^{-1}(1/z))} = z.$$

□

Now, let G_ν^{-1} be the function provided by Theorem 44, then let $R_\nu(z) = G_\nu^{-1}(z) - 1/z$, we have

$$G_\nu(R_\nu(z) + 1/z) = G_\nu(G_\nu^{-1}(z)) = z.$$

Thus, we conclude that for a probability measure ν with variance σ^2, then on the open disc with centre $-i(4\sigma)^{-1}$ and radius $(4\sigma)^{-1}$ there is an analytic function $R_\nu(z)$ such that $G_\nu(R_\nu(z)+1/z) = z$ for $|z+i(4\sigma)^{-1}| < (4\sigma)^{-1}$ where G_ν is the Cauchy transform of ν. Thus, we obtain the R-transform R_ν of the measure ν.

Remark 45. *$G_\nu(R_\nu(z)+1/z) = z$ for $|z+i(4\sigma)^{-1}| < (4\sigma)^{-1}$, if we let $F_\nu(z) = 1/G_\nu(z)$, this becomes $F_\nu(R_\nu(z)+1/z) = 1/z$, and hence $R_\nu(z) = F_\nu^{-1}(1/z) - 1/z$.*

In order to prove next lemma, we deduce some structure in a special Riemann surface. At first, we recall that a *Riemann surface* is a connected, Hausdorff topological manifold X with a countable base for the topology and there exists a family of open sets $\{U_\alpha\}_{\alpha\in I}$ covering X and homeomorphisms $\varphi_\alpha : U_\alpha \to V_\alpha$ where $V_\alpha \subset \mathbb{C}$ is some open set so that

$$\varphi_\alpha \circ \varphi_\beta^{-1} : \varphi_\beta(U_\alpha \cap U_\beta) \to \varphi_\alpha(U_\alpha \cap U_\beta)$$

is holomorphic. We refer to each $(U_\alpha, \varphi_\alpha)$ as a chart.

Let X be a Riemann surface, and $f : X \to \mathbb{C}$ an analytic map. We note that for a fixed $z_0 \in X$, there is a unique $m \geq 0$ and a chart (U, φ) of z_0 such that $\varphi(z_0) = 0$ and $f(\varphi^{-1}(z)) = z^m$ for $z \in \varphi(U)$. We set $mult(f, z_0) = m$. For each $z \in \mathbb{C}$ we define the degree of f at z, denoted $deg_f(z)$, by

$$deg_f(z) = \sum_{w\in f^{-1}(z)} mult(f, w).$$

Theorem 46. ([4]) *If f is proper; that is, the inverse image of a compact set is compact, then deg_f is constant.*

Now. let $F_1, F_2 : \mathbb{C}^+ \to \mathbb{C}^+$ be two analytic functions such that $F_i'(z) \neq 0$ for $z \in \mathbb{C}^+$ and $i = 1, 2$. We set

$$X = \{(z_1, z_2) \in \mathbb{C}^+ \times \mathbb{C}^+ | F_1(z_1) = F_2(z_2)\}.$$

Lemma 47. *X has the structure of a complex manifold so that $(z_1, z_2) \to F_1(z_1)$ is analytic.*

Proof. For $(z_1, z_2) \in X$, we let $w = F_1(z_1) = F_2(z_2)$. For $i = 1, 2$, since $F_i'(z_i) \neq 0$, there is a neighborhood O_i of z_i such that F_i is one-to-one on O_i. F_1 and F_2 are non-constant analytic functions, by Open Mapping Theorem $F_1(O_1)$ and $F_2(O_2)$ are both open. On the other hand, $w \in F_i(O_i)$ for $i = 1, 2$, which implies $F_1(O_1) \cap F_2(O_2)$ is a non-empty open set. We set $U = F_1(O_1) \cap F_2(O_2)$. Then, there are analytic functions f_1, f_2 defined on U such that $F_i \circ f_i$ is identity map on U. We let $V = \{(f_1(u), f_2(u)) | u \in U\}$ and define $\varphi : V \to U$ by $\varphi(w_1, w_2) = F_1(w_1)$.

Now, given two charts (V, φ) and (V', φ'), by the construction above, (z_1, z_2), $(z_1', z_2') \in X$, there are neighborhoods U, U' of $F_1(z_1)$ and $F_1'(z_1')$ and then analytic functions $f_1, f_2 : U \to \mathbb{C}$ and $f_1', f_2' : U' \to \mathbb{C}$ such that $F_i \circ f_i$ =identity, and $F_i \circ f_i'$ =identity. Then, for $u \in \varphi(V \cap V')$, we have $\varphi' \circ \varphi^{-1}(u) = \varphi'(f_1(u), f_2(u)) = F_1(f_1(u)) = u$. So $\varphi' \circ \varphi^{-1}$ is identity, and hence analytic, which implies that X has a structure of a complex manifold. $\square$

Lemma 48. *Suppose F_1,F_2 and X are as in Lemma 47. In addition, we assume there are σ_1, σ_2 such that for $i = 1, 2$, and $z \in \mathbb{C}^+$ we have $|z - F_i(z)| \leq {\sigma_i}^2/Im(z)$. Then $\theta : X \to \mathbb{C}$ given by $\theta(z_1, z_2) = z_1 + z_2 - F_1(z_1)$ is a proper map.*

Proof. Let K be a compact subset of $\mathbb{C}$, we must show that $\varphi^{-1}(K)$ is compact. Without loss of generality, let $K = \overline{B_r(z)}$: the ball with centre z and the radius $r > 0$. We note that $\varphi^{-1}(K)$ is closed since φ is continuous. Hence, it suffices to show that every sequence in $\theta^{-1}(K)$ contains a convergent subsequence. Given a sequence $\{(z_{1,n}, z_{2,n})\}_n$ in $\varphi^{-1}(K)$. Then $-z_{1,n} = -\theta(z_{1,n}, z_{2,n}) + z_{2,n} - F_2(z_{2,n})$, and thus,

$$|z_{1,n}| \leq |\theta(z_{1,n}, z_{2,n})| + |z_{2,n} - F_2(z_{2,n})| \leq |z| + r + \frac{\sigma_2^2}{Im(z_{2,n})}.$$

By Lemma 37, $Im(F_1(z_{1,n})) \geq Im(z_{1,n})$, then

$$Im(z_{2,n}) \geq Im(z_{1,n}) + Im(z_{2,n}) - Im(F_1(z_{1,n})) = Im(\theta(z_{1,n}, z_{2,n})) \geq Im(z) - r.$$

Therefore,

$$|z_{1,n}| \leq |z| + r + \frac{\sigma_2^2}{Im(z) - r}.$$

Similarly, we have

$$|z_{2,n}| \leq |z| + r + \frac{\sigma_1^2}{Im(z) - r}.$$

So there is a subsequence $\{z_{1,n_k}, z_{2,n_k}\}_k$ such that $\{z_{1,n_k}\}$ converges to z_1, and $\{z_{2,n_k}\}$ converges to z_2 for some z_1, z_2. Therefore, $F_1(z_1) = \lim\limits_{k\to\infty} F_1(z_{1,n_k}) = \lim\limits_{k\to\infty} F_2(z_{2,n_{n_k}}) = F_2(z_2)$, which implies that $(z_1, z_2) \in X$. Finally, $\theta(z_1, z_2) = \lim\limits_{k\to\infty} \theta(z_{1,n_l}, z_{2,n_k}) \in K$; hence, $(z_1, z_2) \in \varphi^{-1}(K)$. □

Lemma 49. *Suppose $F_1, F_2 : \mathbb{C}^+ \to \mathbb{C}^+$ are analytic and there is $r > 0$ such that for $z \in \mathbb{C}^+$, and $i = 1, 2$ we have $|F_i(z) - z| \leq r^2/Im(z)$. Then for each $z \in \mathbb{C}^+$, there is a unique pair $(z_1, z_2) \in \mathbb{C}^+ \times \mathbb{C}^+$ such that* (1) *$F_1(z_1) = F_2(z_2)$, and* (2) *$z_1 + z_2 - F_1(z_1) = z$.*

Proof. By Lemma 41, for $i = 1, 2$, we have $1/F_i$ is Cauchy transform of some probability measure, and by Lemma 37 we know that it satisfies $Im(z) \leq Im(F_i(z))$.

First, assume $z \in \mathbb{C}^+_{4r}$, if (z_1, z_2) satisfies (1) and (2), $Im(z_1) = Im(z) + Im(F_2(z_2) - z_2) \geq Im(z)$. Similarly, $Im(z_2) \geq Im(z)$. Hence, (z_1, z_2) should belong to $\mathbb{C}^+_{4r} \times \mathbb{C}^+_{4r}$. By Lemma 43, F_1 and F_2 are invertible on $\mathbb{C}^+_{2r}$. Thus, to find a solution to (1) and (2) is eqnivalent to find $u \in \mathbb{C}^+_{2r}$ such that

$$F_1^{-1}(u) + F_2^{-1}(u) - u = z \tag{4.3}$$

and then let $z_1 = F_1^{-1}(u)$ and $z_2 = F_2^{-1}(u)$. Thus, we must show that for every $z \in \mathbb{C}^+_{4r}$, there is a unique $u \in \mathbb{C}^+_{2r}$ satisfies equation (4.3). Let C be the circle with centre z and radius $2r$. Then $C \subset \mathbb{C}^+_{2r}$ and for $u \in C$, by Lemma 43 we have

$$|F_1^{-1}(u) - u| + |F_2^{-1}(u) - u| \leq \frac{4r^2}{Im(u)} < \frac{4r^2}{2r} = 2r.$$

Thus,

$$|(z-u)-[z-u-(F_1^{-1}(u)-u)-(F_2^{-1}(u)-u)]| \leq 2r = |z-u|.$$

Therefore by Rouché's Theorem, there is a unique u belong to the inside of C such that

$$z-u=(F_1^{-1}(u)-u)+(F_2^{-1}(u)-(u)).$$

Suppose that there is $v \in \mathbb{C}_{2r}^+$ with

$$z-v=(F_1^{-1}(v)-v)+(F_2^{-1}(v)-v).$$

Then, we note that

$$|z-v| = |(F_1^{-1}(v)-v)+(F_2^{-1}(v)-v)| < 2r,$$

and hence v belong to the inside of C. So, there is a unique $u \in \mathbb{C}_{2r}^+$ satisfying equation (4.3).

Let $X = \{(z_1,z_2)|F_1(z_1)=F_2(z_2)\}$. By Lemma 47, X is a Riemann surface. Let $\theta(z_1,z_2)=z_1+z_2-F_1(z_1)$. We have showed that for $z \in \mathbb{C}_{4r}^+$, $deg_\theta = 1$. On the other hand, by Lemma 48 and Theorem 46, deg_θ is constant on $\mathbb{C}^+$. So there is a unique solution to (1) and (2) for all $\mathbb{C}^+$. □

Remark 50. *Let ν be a probability measure with σ^2 and mean m. Let $\widetilde{\nu}(E)=\nu(E+m)$, then $\widetilde{\nu}$ is a pribability measure with mean zero and variance σ^2. If G_ν and $\widetilde{G_{\widetilde{\nu}}}$ be the corresponding Cauchy transforms, then*

$$\widetilde{G}_{\widetilde{\nu}}(z)=\int_{\mathbb{R}}\frac{1}{z-t}d\widetilde{\nu}(t)=\int_{\mathbb{R}}\frac{1}{z-(t-m)}d\nu(t)=\int_{\mathbb{R}}\frac{1}{(z+m)-t}d\nu(t)=G_\nu(z+m).$$

Thus, if we let $\widetilde{F_{\widetilde{\nu}}}=1/\widetilde{G}_{\widetilde{\nu}}$, by Lemma 38 there is a probability measure ρ such that

$$z-\widetilde{F_{\widetilde{\nu}}}(z)=\sigma^2\int_{\mathbb{R}}\frac{1}{z-t}d\rho(t).$$

Finally, let $\widetilde{R_{\widetilde{\nu}}}$ be the R-transform of $\widetilde{\nu}$, then $G_\nu(\widetilde{R_{\widetilde{\nu}}}(z)+m+1/z)=\widetilde{G}_{\widetilde{\nu}}(\widetilde{R_{\widetilde{\nu}}}(z)+1/z)=z$ for $|z+i(4\sigma)^{-1}|<(4\sigma)^{-1}$. Thus, $R_\nu(z)=\widetilde{R_{\widetilde{\nu}}}(z)+m$ for $|z+i(4\sigma)^{-1}|<(4\sigma)^{-1}$.

Theorem 51. *Let μ,ν be probability measures on $\mathbb{R}$ with finite variances and R_μ, R_ν be the corresponding R-transforms. Then there is a unique probability measure with finite variance, denoted $\mu \boxplus \nu$. and called the free convolution of μ and ν, such that*

$$R_{\mu\boxplus\nu}=R_\mu+R_\nu,$$

where $R_{\mu\boxplus\nu}$ is the R-transform of $\mu\boxplus\nu$. Furthermore, the first moment of $\mu\boxplus\nu$ is the sum of the first moments of μ and ν and the variance of $\mu\boxplus\nu$ is the sum of the variances of μ and ν.

Proof. By Remark 50, it suffices to show the theorem in the case μ and ν are centered. For let $\sigma_\mu^2, \sigma_\nu^2$ be the variance of the probability measure μ and ν. By Lemma 38 there are probability measures ρ_1 and ρ_2 such that for $z \in \mathbb{C}^+$, we have

$$z-F_\mu(z)=\sigma_\mu^2\int_{\mathbb{R}}\frac{d\rho_1(t)}{z-t} \text{ and } z-F_\nu(z)=\sigma_\nu^2\int_{\mathbb{R}}\frac{d\rho_2(t)}{z-t}$$

where $F\mu = 1/G_\mu$, $F\nu = 1/G_\nu$, and G_μ, G_ν are the Cauchy transforms of μ, ν.

By Lemma 49, for each $z \in \mathbb{C}^+$ there is a unique pair $(z_1, z_2) \in \mathbb{C}^+ \times \mathbb{C}^+$ such that $F_\mu(z_1) = F_\nu(z_2)$ and $z_1 + z_2 - F_\mu(z_1) = z$. Define $F(z) = F_\mu(z_1)$. Now, let $X = \{(z_1, z_2) | F_\mu(z_1) = F_\nu(z_2)\}$ and $\theta : X \to \mathbb{C}^+$ be as in Lemma 48, then θ is analytic and bijection since $deg_\theta = 1$. We define $\pi(z_1, z_2) = z_\mu$, and then $F = F_\mu \circ \pi \circ \theta^{-1}$ is analytic on $\mathbb{C}^+$ and we have

$$z - F(z) = z_1 - F_\mu(z_1) + z_2 - F_\nu(z_2). \tag{4.4}$$

Since $Im(F_\mu(z_1)) \geq Im(z_1)$,

$$\begin{aligned} z = F(z) + z_1 - F_\mu(z_1) + z_2 - F_\nu(z_2) \quad &\Rightarrow \quad z = z_2 + z_1 - F_\mu(z_1) \\ &\Rightarrow \quad Im(z) = Im(z_2) + Im(z_1 - F_\mu(z_1)) \leq Im(z_2). \end{aligned}$$

Similarly, $Im(z) \leq Im(z_1)$. Hence, we have

$$|z - F(z)| = |z_1 - F_\mu(z_1) + z_2 - F_\nu(z_2)| \leq \frac{\sigma_\mu^2}{Im(z_1)} + \frac{\sigma_\nu^2}{Im(z_2)} \leq \frac{\sigma_\mu^2 + \sigma_\nu^2}{Im(z)}.$$

Thus, by Lemma 41, $1/F$ is the Cauchy transform of a centered probability measure with variance $\sigma^2 \leq \sigma_\mu^2 + \sigma_\nu^2$. Thus, by Lemma 35, there is a probability measure ρ such that

$$z - F(z) = \sigma^2 \int_{\mathbb{R}} \frac{1}{z - t} d\rho(t).$$

By equation (4.4), we have

$$\sigma^2 \int_{\mathbb{R}} \frac{1}{z - t} d\rho(t) = \sigma_\mu^2 \int_{\mathbb{R}} \frac{1}{z_1 - t} d\rho_1(t) + \sigma_\nu^2 \int_{\mathbb{R}} \frac{1}{z_2 - t} d\rho_2(t)$$

and therefore

$$\sigma^2 \int_{\mathbb{R}} \frac{z}{z - t} d\rho(t) = \sigma_\mu^2 \int_{\mathbb{R}} \frac{z}{z_1 - t} d\rho_1(t) + \sigma_\nu^2 \int_{\mathbb{R}} \frac{z}{z_2 - t} d\rho_2(t). \tag{4.5}$$

For $z = iy,\ y > 0$, by Lemma 40 we have $|iy - t| \geq y/\sqrt{1 + \alpha^2}$ for any $\alpha > 0$. Then

$$|1 - \frac{1}{z} F(z)| = |\frac{1}{z}||F(z) - z| = \frac{\sigma^2}{y} |\int_{\mathbb{R}} \frac{1}{iy - t} d\rho(t)| \leq \frac{\sigma^2}{y} \int_{\mathbb{R}} \frac{1}{|iy - t|} d\rho(t) \leq \frac{\sigma^2 \sqrt{1 + \alpha^2}}{y^2}.$$

Therefore, for $z = iy,\ y > 0$, we have $F(z)/z \to 1$. Now, by Lemma 43 part(3) and (4), we have

$$|F_\mu^{-1}(F(z)) - F(z)| \leq \frac{2\sigma^2}{Im(F(z))} \leq \frac{2\sigma^2}{Im(z)}.$$

We note that

$$\frac{z_1}{z} = \frac{F_\mu^{-1}(F(z)) - F(z)}{z} + \frac{F(z)}{z}.$$

Then, if $z = iy,\ y > 0$, the first term goes to 0, and the second term goes to 1 as $y \to \infty$. Hence, $z_1/(iy) \to 1$ as $y \to \infty$. Similarly, $z_2/(iy) \to 1$ as $y \to \infty$. Now, by applying Lemma 40, we have

$$|\frac{iy}{z_1 - t}| \leq \frac{y}{|z_1 - t|} \leq \frac{1}{\sqrt{1 + \alpha^2}} \frac{y}{|z_1|} < M \quad \text{for some } M > 0 \text{ since } z_1/(iy) \to 1 \text{ as } y \ \to \infty.$$

Thus, by Dominated Convergence Theorem,

$$\lim_{y\to\infty}\int_{\mathbb{R}} \frac{iy}{z_1 - t} d\rho_1(t) = 1 \quad \text{and likewise} \quad \lim_{y\to\infty}\int_{\mathbb{R}} \frac{iy}{z_2 - t} d\rho_2(t) = 1.$$

Now, if we take limits as $y \to \infty$ in equation (4.5), the left hand side will goes to σ^2 by Lemma 40(4), and the right hand side will goes to $\sigma_\mu^2 + \sigma_\nu^2$. Hence, we have $\sigma^2 = \sigma_\mu^2 + \sigma_\nu^2$.

Let $D_r = \{z \mid |z + ir| < r\}$. Since σ is greater than $\sigma_\mu, \sigma_\nu > 0$ and note that $|z + i(4\sigma)^{-1}| < (4\sigma)^{-1}$ if and only if $Im(1/z) > 2\sigma$, we have $D_{1/(4\sigma)} \subset D_{1/(4\sigma_\mu)} \cap D_{1/(4\sigma_\nu)}$.

Thus, for $z \in D_{1/(4\sigma)}$, then $1/z$ is in the domains of F^{-1}, F_μ^{-1}, and F_ν^{-1}. Now for $F^{-1}(1/z)$, by Lemma 45 there are $z_1, z_2 \in \mathbb{C}^+$ such that $F_\mu(z_1) = F_\nu(z_2)$ and $F^{-1}(1/z) = z_1 + z_2 - F_\mu(z_1)$. However, by the construction of F we have

$$1/z = F(F^{-1}(1/z)) = F_\mu(z_1) = F_\nu(z_2) \text{ and hence } z_1 = F_\mu^{-1}(1/z),\ z_2 = F_\nu^{-1}(1/z).$$

Thus,

$$\begin{aligned} & F^{-1}(1/z) = z_1 + z_2 - F_\mu(z_1) = F_\mu^{-1}(1/z) + F_\nu^{-1}(1/z) - 1/z \\ \Rightarrow\quad & F^{-1}(1/z) - 1/z = (F_\mu^{-1}(1/z) - 1/z) + (F_\nu^{-1}(1/z) - 1/z). \end{aligned}$$

Finally, by Remark 45 we conclude that $R(z) = R_\mu(z) + R_\nu(z)$. $\square$

Chapter 5

Free Convolution: General Measures

In this chapter, we consider general probability measures on $\mathbb{R}$. Our final goal is to prove Theorem 69. That is, if we have two general probability measures μ and ν on $\mathbb{R}$ and R_μ, R_ν be the R-transforms of μ and ν, then there is a probability measure $\mu \boxplus \nu$ such that

$$R_{\mu\boxplus\nu} = R_\mu + R_\nu,$$

where $R_{\mu\boxplus\nu}$ is the R-transform of $\mu \boxplus \nu$.

For $\alpha, \beta > 0$, we consider the following notation

$$\triangle_{\alpha,\beta} = \{1/z | z \in \Gamma_{\alpha,\beta}\} = \{w \in \mathbb{C} | -\alpha Im(w) < Re(w) \text{ and } |w + \frac{i}{2\beta}| < \frac{1}{2\beta}\} \subset \mathbb{C}^-.$$

Lemma 52. *Let ν be a probability measure on $\mathbb{R}$ and G_ν its Cauchy transform. Let $F_\nu = 1/G_\nu$. Suppose $0 < \alpha_1 < \alpha_2$. Then there is $\beta_0 > 0$ such that for all $\beta_2 \geq \beta_0$ and $\beta_1 \geq \beta_2(1 + \alpha_2 - \alpha_1)$, we have $\Gamma_{\alpha_1,\beta_1} \subseteq F_\nu(\Gamma_{\alpha_2,\beta_2})$ and for each $w \in \Gamma_{\alpha_1,\beta_1}$ there is a unique $z \in \Gamma_{\alpha_2,\beta_2}$ such that $F_\nu(z) = w$. Hence, there is an analytic function F_ν^{-1}, defined on $\Gamma_{\alpha_1,\beta_1}$ such that $F_\nu(F_\nu^{-1}(w)) = w$.*

Proof. Let $\theta = \tan^{-1}(\alpha_1^{-1}) - \tan^{-1}(\alpha_2^{-1})$. Then

$$\begin{aligned} \sin(\theta) &= \sin(\tan^{-1}(\alpha_1^{-1}) - \tan^{-1}(\alpha_2^{-1})) \\ &= \sin(\tan^{-1}(\alpha_1^{-1}))\cos(\tan^{-1}(\alpha_2^{-1})) - \cos(\tan^{-1}(\alpha_1^{-1}))\sin(\tan^{-1}(\alpha_2^{-1})) \\ &= \frac{\alpha_2 - \alpha_1}{\sqrt{1+\alpha_1^2}\sqrt{1+\alpha_2^2}}. \end{aligned}$$

Choose $\varepsilon > 0$ such that $\varepsilon < \sin(\theta)$. By Lemma 40(4), we choose $\beta_0 > 0$ such that $|F_\nu(z) - z| < \varepsilon|z|$ for all $z \in \Gamma_{\alpha_2,\beta_0}$. Let $\beta_2 \geq \beta_0$ and $\beta_1 \geq \beta_2(1 + \alpha_2 - \alpha_1)$.

Firstly, we show that for $w \in \Gamma_{\alpha_1,\beta_1}$ and for $z \in \partial\Gamma_{\alpha_2,\beta_2}$, we have $\varepsilon|z| < |z - w|$. If $z = \alpha_2 y + iy \in \partial\Gamma_{\alpha_2}$, then $|z - w|/|z| \geq \sin(\theta) > \varepsilon$. If $z = x + i\beta_2 \in \partial\Gamma_{\alpha_2,\beta_2}$, then

$$\begin{aligned} |z - w| > \beta_1 - \beta_2 \geq \beta_2(\alpha_2 - \alpha_1) &= \beta_2 \sin(\theta)\sqrt{1+\alpha_1^2}\sqrt{1+\alpha_2^2} \\ &> \varepsilon\beta_2\sqrt{1+\alpha_1^2}\sqrt{1+\alpha_2^2} \geq \varepsilon|z|\sqrt{1+\alpha_1^2} > \varepsilon|z|. \end{aligned}$$

Then, for $w \in \Gamma_{\alpha_1,\beta_1}$ and $z \in \partial\Gamma_{\alpha_2,\beta_2}$ we have $\varepsilon|z| < |z - w|$.

Now, fix $w \in \Gamma_{\alpha_1,\beta_1}$ and let $r > |w|/(1 - \varepsilon)$. Thus, for $z \in \{\tilde{z} | |\tilde{z}| = r\} \cap \Gamma_{\alpha_2,\beta_2}$ we have

$$|z - w| \geq |z| - |w| = r - |w| > \varepsilon r = \varepsilon|z|.$$

Now, we let

$$C = (\partial\Gamma_{\alpha_2,\beta_2} \cap \{\tilde{z}||\tilde{z}| \leq r\}) \cup (\{\tilde{z}||\tilde{z}| = r\} \cap \Gamma_{\alpha_2,\beta_2}).$$

Then, C is the curve such that for $z \in C$, we have that $\varepsilon|z| < |z-w|$. Thus, for $z \in C$, we have

$$|(F_\nu(z) - w) - (z - w)| < \varepsilon|z| < |z-w|.$$

Thus, by Rouché's Theorem, there is a unique z in the interior of C such that $F_\nu(z) = w$. Since r is arbitrary chosen such that $r > |w|/(1-\varepsilon)$, we can let r as big as we want. Therefore, we conclude that there is a unique $z \in \Gamma_{\alpha_2,\beta_2}$ such that $F_\nu(z) = w$.

On the other hand, since F_ν is analytic and one-to-one on the inverse image of $\Gamma_{\alpha_1,\beta_1}$, the inverse function F_ν^{-1} of F_ν is analytic. □

Lemma 53. *Let $F_\nu = 1/G_\nu$ where G_ν is the Cauchy transform of a probability measure ν on $\mathbb{R}$. Suppose $0 < \alpha_1 < \alpha_2$. Then there is $\beta_0 > 0$ such that $F_\nu(\Gamma_{\alpha_1,\beta_1}) \subseteq \Gamma_{\alpha_2,\beta_1}$ for all $\beta_1 \geq \beta_0$.*

Proof. Choose $0 < \varepsilon < 1/2$ so that

$$\alpha_1 < \frac{\alpha_1 + \varepsilon/\sqrt{1-\varepsilon^2}}{1 - \alpha_1\varepsilon/\sqrt{1-\varepsilon^2}} < \alpha_2.$$

Then, by Lemma 40(4) we can choose $\beta_0 > 0$ such that $|F_\nu(z) - z| < \varepsilon/2|z|$ for $z \in \Gamma_{\alpha_1,\beta_0}$. Observe that for such z,

$$1 - \frac{|F_\nu(z)|}{|z|} \leq |\frac{F_\nu(z)}{z} - 1| < \frac{\varepsilon}{2} \Rightarrow \frac{|F_\nu(z)|}{|z|} > 1 - \frac{\varepsilon}{2} \Rightarrow \frac{|z|}{|F_\nu(z)|} < \frac{1}{1-\varepsilon/2} < \frac{4}{3}.$$

Suppose $\beta_1 \geq \beta_0$ and let $z \in \Gamma_{\alpha_1,\beta_1}$. We assume that $Re(z) \geq 0$ (the case $Re(z) < 0$ is similar). Write $z = |z|e^{i\phi}$. Then, $\phi > \tan^{-1}(\alpha_1^{-1})$. Write $F_\nu(z) = |F_\nu(z)|e^{i\psi}$,

$$|F_\nu(z) - z| < \frac{\varepsilon}{2} \Rightarrow |\frac{|F_\nu(z)|}{|z|}e^{i(\psi-\phi)} - 1| < \frac{\varepsilon}{2} \Rightarrow |\frac{|F_\nu(z)|}{|z|}\sin(\psi - \phi)| < \frac{\varepsilon}{2} \Rightarrow |\sin(\psi-\phi)| < \frac{\varepsilon}{2}\frac{|z|}{|F_\nu(z)|} < \frac{4\varepsilon}{6}.$$

Thus $|\sin(\psi - \phi)| < \varepsilon$, so

$$\psi > \phi - \sin^{-1}(\varepsilon) > \tan^{-1}(\alpha_1^{-1}) - \sin^{-1}(\varepsilon).$$

If $\psi \leq \pi/2$, then

$$\begin{aligned}\tan(\psi) > \tan(\tan^{-1}(\alpha_1^{-1}) - \sin^{-1}(\varepsilon)) &= \frac{\tan(\tan^{-1}(\alpha_1^{-1})) - \tan(\sin^{-1}(\varepsilon))}{1 + \tan(\tan^{-1}(\alpha_1^{-1}))\tan(\sin^{-1}(\varepsilon))} \\ &= \frac{\alpha_1^{-1} - \varepsilon/\sqrt{1-\varepsilon^2}}{1 + \alpha_1^{-1}\varepsilon\sqrt{1-\varepsilon^2}} = \frac{1 - \alpha_1\varepsilon/\sqrt{1-\varepsilon^2}}{\alpha_1 + \varepsilon/\sqrt{1-\varepsilon^2}} > \frac{1}{\alpha_2}.\end{aligned}$$

Then, $F_\nu(z) \in \Gamma_{\alpha_2}$. Suppose $\psi \geq \pi/2$. Then we claim $\pi - \psi > \tan^{-1}(\alpha_2^{-1})$. Since $|\psi - \phi| < \sin^{-1}(\varepsilon)$ and $\phi \leq \pi/2$ ($Re(z) \geq 0$), we have $\psi - \phi < \sin^{-1}(\varepsilon)$, which implies that $\psi < \sin^{-1}(\varepsilon) + \phi \leq \sin^{-1}(\varepsilon) + \pi/2$. Therefore, $\pi - \psi > \pi/2 - \sin^{-1}(\varepsilon)$. Thus,

$$\tan(\pi - \psi) > \tan(\pi/2 - \sin^{-1}(\varepsilon)) = \frac{\sqrt{1-\varepsilon^2}}{\varepsilon}.$$

On the other hand,

$$\alpha_2 > \frac{\alpha_1 + \varepsilon/\sqrt{1-\varepsilon^2}}{1 - \alpha_1\varepsilon/\sqrt{1-\varepsilon^2}} > \alpha_1 + \frac{\varepsilon}{\sqrt{1-\varepsilon^2}} > \frac{\varepsilon}{\sqrt{1-\varepsilon^2}},$$

so $\tan(\pi-\psi) > \alpha_2^{-1}$, and hence $\pi - \psi > \tan^{-1}(\alpha_2^{-1})$. Thus, $F_\nu(z) \in \Gamma_{\alpha_2}$.

Since we also have $Im(F_\nu(z)) \geq Im(z) > \beta_1$ by Lemma 37, we obtain $F_\nu(z) \subseteq \Gamma_{\alpha_2,\beta_1}$. □

Theorem 54. *Let ν be a probability measure on $\mathbb{R}$ with Cauchy transform G_ν and set $F_\nu = 1/G_\nu$. For every $\alpha > 0$ there is $\beta > 0$ such that $R_\nu(z) = F_\nu^{-1}(1/z) - 1/z$ is defined for $z \in \triangle_{\alpha,\beta}$ and R_ν satisfies $G_\nu(R_\nu(z) + 1/z) = z$ for $z \in \triangle_{\alpha,\beta}$ and $R_\nu(G_\nu(z)) + 1/G_\nu(z) = z$ for $z \in \Gamma_{\alpha,\beta}$.*

Proof. Let $F_\nu(z) = 1/G_\nu(z)$. Given $\alpha > 0$, Lemma 52 implies that we can choose $\beta_0 > 0$ such that F_ν^{-1} is defined on $\Gamma_{2\alpha,\beta_0}$. Thus, for $z \in \triangle_{2\alpha,\beta_0}$, we define $R_\nu(z) = F_\nu^{-1}(1/z) - 1/z$. Moreover,

$$G_\nu(R_\nu(z) + 1/z) = G_\nu(F_\nu^{-1}(1/z)) = z.$$

By Lemma 53, there is $\beta > \beta_0$ such that $F_\nu(\Gamma_{\alpha,\beta}) \subseteq \Gamma_{2\alpha,\beta}$. Now, for $z \in \Gamma_{\alpha,\beta}$, we have $G_\nu(z) = 1/F_\nu(z) \in \triangle_{2\alpha,\beta} \subseteq \triangle_{2\alpha,\beta_0}$. Thus,

$$R_\nu(G_\nu(z)) + \frac{1}{G_\nu(z)} = F_\nu^{-1}(\frac{1}{G_\nu(z)}) - \frac{1}{G_\nu(z)} + \frac{1}{G_\nu(z)} = F_\nu^{-1}(F_\nu(z)) = z.$$

However, $\triangle_{\alpha,\beta} \subseteq \triangle_{2\alpha,\beta_0}$, which implies that $G_\nu(R_\nu(z) + 1/z) = z$ for $z \in \triangle_{\alpha,\beta}$. □

According to Theorem 54, we define the R-transform of a general measure on $\mathbb{R}$ as follows:

Definition 55. Let ν be a probability measure on $\mathbb{R}$, let G_ν be the Cauchy transform of ν. We define the R-transform of ν as the germ of analytic functions on $\triangle_{\alpha,\beta}$ satisfying the equation

$$G_\nu(R_\nu(z) + 1/z) = z = \frac{1}{G_\nu(z)} + R_\nu(G_\nu(z)).$$

That is, for all $\alpha > 0$ there is $\beta > 0$ such that $G_\nu(R_\nu(z)+1/z) = z$ for all $z \in \triangle_{\alpha,\beta}$ and $1/G_\nu(z)+R_\nu(G_\nu(z)) = z$. for all $z \in \Gamma_{\alpha,\beta}$.

Remark 56. *In the definition 55 we say that R_ν is a germ of analytic functions means that for $i = 1, 2$, given $\alpha_i > 0$ there is $\beta_i > 0$ and an analytic function $R_\nu^{(i)}$ defined on $\triangle_{\alpha_i,\beta_i}$ such that $R_\nu^{(1)} = R_\nu^{(2)}$ on $\triangle_{\alpha_1,\beta_1} \cap \triangle_{\alpha_2,\beta_2}$.*

Remark 57. *For the remaining part of this chapter, we shall show that the R-transform addition formula still hold. That is, given probability measures μ and ν with R-transform R_μ and R_ν respectively, there is a probability measure m with Cauchy transform G_m and R-transform R_m such that $R_m = R_\mu + R_\nu$. This means that for all $\alpha > 0$ there is $\beta > 0$ such that R_m, R_μ and R_ν are defined on $\triangle_{\alpha,\beta}$ and for $z \in \triangle_{\alpha,\beta}$, we have $R_m(z) = R_\mu(z) + R_\nu(z)$. We shall denote m by $\mu \boxplus \nu$ and call it the free convolution of μ and ν.*

We note that if μ is a Dirac mass at $a \in \mathbb{R}$, then $G_\mu(z) = (z-a)^{-1}$, and then $R_\mu(z) = a$. So $R_m(z) = a + R_\nu(z)$ and thus $G_m(z) = G_\nu(z+a)$. Therefore, $m(E) = \nu(E-a)$ where E is a Borel set in $\mathbb{R}$. As a result, for the rest of this chapter, we assume that both μ and ν are not Dirac masses.

Now, we consider another case: given $a+bi \in \mathbb{C}$ such that $b<0$ and let μ be the probability measure with density $d\mu(t) = (1/\pi)(-b/b^2+(t-a)^2)dt$. Then,

$$G_\mu(z) = \int_{\mathbb{R}} \frac{1}{z-t}\frac{1}{\pi}\frac{-b}{b^2+(t-a)^2}dt = \frac{b}{\pi}\int_{\mathbb{R}} \frac{1}{(t-z)(t-a-bi)(t-a+bi)}dt.$$

We let $f(w) = [(w-z)(w-a-bi)(w-a+bi)]^{-1}$ where $w \in \mathbb{C}$ and let Γ_R be the closed curve formed by joining part of the circle $|w|=R$ in $\mathbb{C}^+$ to the interval $[-R,R]$. Note that

$$G_\mu(z) = \frac{b}{\pi}\lim_{R\to\infty}\int_{\Gamma_R} f(w)dw.$$

It is clear that $a-bi$ and z will be in the interior of the circle when R is large, and $Res(f, a-bi) = \lim_{w\to(a-bi)}[w-(a-bi)]f(w) = [2bi(z-(a-bi))]^{-1}$, and $Res(f,z) = \lim_{w\to z}(w-z)f(w) = [(z-(a+bi))(z-(a-bi))]^{-1}$. Hence, by Residue Theorem,

$$G_\mu(z) = \frac{b}{\pi}\lim_{R\to\infty} 2\pi i(Res(f,a-bi)+Res(f,z)) = \frac{1}{z-(a+bi)} = \frac{1}{z-w}.$$

Hence, $G_\mu^{-1}(z) = 1/z - w$, which implies that the R-transform $R_\mu(z) = w$. Let ν be any probability measure on $\mathbb{R}$, and let G_ν be the Cauchy transform of ν and R_ν be its R-transform. So if $\mu \boxplus \nu$ exists, its R-transform should be $R_{\mu\boxplus\nu}(z) = w + R_\nu(z)$ and we denote its Cauchy transform by $G_{\mu\boxplus\nu}$. Then, we have $z = G_{\mu\boxplus\nu}(R_{\mu\boxplus\nu}(z)+1/z) = G_{\mu\boxplus\nu}(R_\mu(z)+R_\nu(z)+1/z)$. If we put $s = R_{\mu\boxplus\nu}(z)+1/z$, then we have $G_{\mu\boxplus\nu}(s) = z$ and

$$G_{\mu\boxplus\nu}(s) = z = G_\nu(R_\nu(z)+1/z) = G_\nu(s-R_\mu(z)) = G_\nu(s-R_\mu(G_{\mu\boxplus\nu}(s))).$$

If we define the subordination functions $\omega_1(z) = z - R_\nu(G_{\mu\boxplus\nu}(z))$, and $\omega_2(z) = z - R_\mu(G_{\mu\boxplus\nu}(z))$. Then, $\omega_2(z) = z-w$ and

$$G_{\mu\boxplus\nu}(z) = G_\nu(z-R_\mu(G_{\mu\boxplus\nu}(z))) = G_\nu(\omega_2(z)).$$

Now consider $\omega_2(z) = z-w$ is a maps $\mathbb{C}^+$ to $\mathbb{C}^+$ and letting $G = G_{\nu_2}\circ\omega_2$, we have $\lim\limits_{y\to\infty} iyG_{\mu\boxplus\nu}(iy) = 1$ by Proposition 26. By Theorem 27, there is a measure, which we denote it by $\mu\boxplus\nu$, of which $G_{\mu\boxplus\nu}$ is the Cauchy transform and thus the R-transform of this measure satisfies the equation $R_{\mu\boxplus\nu} = R_\mu + R_\nu$ by construction.

By Theorem 28, we have $\mu\boxplus\nu = \mu * \nu$ where $*$ means the classical convolution, because $G_{\mu\boxplus\nu}(z) = G_\nu(z-w) = G_{\mu*\nu}(z)$. We summarize previous discussion by the following theorem:

Theorem 58. *Given $a+bi \in \mathbb{C}^-$. Let μ_0 be the Dirac mass at a and μ_1 be the probability measure with density*

$$d\mu_1(t) = \frac{1}{\pi}\frac{-b}{b^2+(t-a)^2}dt.$$

Then for any probability measure ν, we have $\mu_i \boxplus \nu = \mu_i * \nu$ for $i = 0,1$.

Now, given two probability measures μ and ν on $\mathbb{R}$, observe that if $\mu\boxplus\nu$ exists and the R-transform addition formula hold. That is, $R_{\mu\boxplus\nu} = R_\mu + R_\nu$. Then,

$$G_{\mu\boxplus\nu}^{-1}(z) = G_\mu^{-1}(z) + G_\nu^{-1}(z) - 1/z \Longleftrightarrow z = G_\mu^{-1}(G_{\mu\boxplus\nu})(z) + G_\nu^{-1}(G_{\mu\boxplus\nu}(z)) - \frac{1}{G_{\mu\boxplus\nu}(z)}$$

Now, if we let $\omega_1(z) = G_\mu^{-1}(G_{\mu\boxplus\nu})(z)$ and $\omega_2(z) = G_\nu^{-1}(G_{\mu\boxplus\nu}(z))$, then

$$z = \omega_1(z) + \omega_2(z) - F_{\mu\boxplus\nu}(z),$$

where $F_{\mu\boxplus\nu} = 1/G_{\mu\boxplus\nu}$. On the other hand,

$$\omega_1(z) = G_\mu^{-1}(G_{\mu\boxplus\nu}(z)) = F_\mu^{-1}(F_{\mu\boxplus\nu}(z)) \Rightarrow F_{\mu\boxplus\nu}(z) = F_\mu(\omega_1(z)).$$

Similarly, $F_{\mu\boxplus\nu}(z) = F_\nu(\omega_2(z))$. Thus, in order to define $\mu \boxplus \nu$ in full generality, we are going to show that we can always find ω_1 and ω_2 satisfying

$$F_\mu(\omega_1(z)) = F_\nu(\omega_2(z)) \text{ and } \omega_1(z) + \omega_2(z) = z + F_\mu^{-1}(\omega_1(z)).$$

Notation 59. *Let μ and ν be probability measures on $\mathbb{R}$ with Cauchy transforms G_μ and G_ν respectively. Let $F_\mu(z) = 1/G_\mu(z)$, $F_\nu(z) = 1/G_\nu(z)$ and $H_\mu(z) = F_\mu(z) - z$, $H_\nu(z) = F_\nu(z) - z$.*

The functions F_μ, F_ν are analytic functions that map $\mathbb{C}^+$ into $\mathbb{C}^+$ obvious. In addition, by Lemma 37 $Im(F_\mu(z))$ and $Im(F_\nu(z))$ are greater than $Im(z)$, we also have H_μ and H_ν are analytic functions that map the upper half plane $\mathbb{C}^+$ to itself.

Corollary 60. *Let F_μ and F_ν be as in Notation 59. Suppose $0 < \alpha_2 < \alpha_1$. Then there are $\beta_2 \geq \beta_0 > 0$ such that*
(1) F_μ^{-1} is defined on $\Gamma_{\alpha_1,\beta_1}$ for any $\beta_1 \geq \beta_0$ with $F_\mu^{-1}(\Gamma_{\alpha_1,\beta_1}) \subseteq \Gamma_{\alpha_1+1,\beta_1/2}$;
(2) $F_\nu(\Gamma_{\alpha_2,\beta_2}) \subseteq \Gamma_{\alpha_1,\beta_0}$.

Proof. Let $\alpha = \alpha_1 + 1$. By Lemma 52, there is $\beta_0/2 > 0$ such that for all $\beta \geq \beta_0/2$ and $\beta_1 = \beta(1+\alpha+\alpha_1) = 2\beta \geq \beta_0$ we have $\Gamma_{\alpha_1,\beta_1} \subseteq F_1(\Gamma_{\alpha,\beta})$ and F_μ^{-1} exists on $\Gamma_{\alpha_1,\beta_1}$; thus, $F_\mu^{-1}(\Gamma_{\alpha_1,\beta_1}) \subseteq \Gamma_{\alpha_1+1,\beta_1/2}$. By Lemma 53, choose $\beta_2 \geq \beta_0 > 0$ so that $F_\nu(\Gamma_{\alpha_2,\beta_2}) \subseteq \Gamma_{\alpha_1,\beta_2} \subseteq \Gamma_{\alpha_1,\beta_0}$. $\square$

Definition 61. For any $z, w \in \mathbb{C}^+$, let $g(z,w) = z + H_\mu(z + H_\nu(w))$. Then $g : \mathbb{C}^+ \times \mathbb{C}^+ \to \mathbb{C}^+$ is analytic. Let $g_z(w) = g(z,w)$.

Remark 62. *Choose $\alpha_1 > \alpha_2 > 0$, and $\beta_2 \geq \beta_0 \geq 0$ according to Corollary 60. Now, we try to control $Im(F_\nu(w) - w)$. Since $F_\nu(w)/w \to 1$ as $w \to \infty$ in $\Gamma_{\alpha_2,\beta_2}$, we have for any $\varepsilon < 1$, $|F_\nu(w) - w| < \varepsilon|w|$ for sufficiently large $w \in \Gamma_{\alpha_2,\beta_2}$. Then*

$$0 \leq Im(F_\nu(w) - w) \leq |F_\nu(w) - w| < \varepsilon|w| < \varepsilon\sqrt{1+\alpha_2^2}\,Im(w);$$

the latter inequality is from Lemma 37 for $w \in \Gamma_{\alpha_2,\beta_2}$. By choosing $1/\varepsilon = 2\sqrt{1+2\alpha_2^2}$ we find a $\beta > 0$ with $\beta \geq \beta_2$ such that we have

$$Im(F_\nu(w) - w) < \frac{1}{2}Im(w) \quad \textit{for all } w \in \Gamma_{\alpha_2,\beta} \subseteq \Gamma_{\alpha_2,\beta_2}.$$

Consider the point $z = w + F_\mu^{-1}(F_\nu(w)) - F_\nu(w)$ where $w \in \Gamma_{\alpha_2,\beta}$. Since $F_\nu(w) \in \Gamma_{\alpha_1,\beta_0}$, this is well-defined. Moreover, we have $Im(F_\nu(w)) \geq Im(w) > \beta \geq \beta_0$, and thus $F_\nu(w) \in \Gamma_{\alpha_1,Im(w)}$, which also implies that $F_\mu^{-1}(F_\nu(w)) \in \Gamma_{\alpha_1+1,Im(w)/2}$ due to Corollary 56(1). Hence, $Im(F_\mu^{-1}(F_\nu(w))) > Im(w)/2$, and therefore

$$Im(z) = Im(F_\mu^{-1}(F_\nu(w))) - Im(F_\nu(w) - w) > \frac{Im(w)}{2} - \frac{Im(w)}{2} = 0.$$

We conclude that for $w \in \Gamma_{\alpha_2,\beta}$, we have $z \in \mathbb{C}^+$.

Lemma 63. *Consider α_2 and β as above, let $w \in \Gamma_{\alpha_2,\beta}$. Then*

$$z = w + F_\mu^{-1}(F_\nu(w)) - F_\nu(w) \Longleftrightarrow g(z, w) = w.$$

Proof. Suppose $z = w + F_\mu^{-1}(F_\nu(w)) - F_\nu(w)$. By Remark 62, we have $z \in \mathbb{C}^+$. Then

$$\begin{aligned} g(z,w) = z + H_\mu(z + H_\nu(w)) &= z + H_\mu(z + F_\nu(w) - w) = z + H_\mu(F_\mu^{-1}(F_\nu(w))) \\ &= z + F_\mu(F_\mu^{-1}(F_\nu(w))) - F_\mu^{-1}(F_\nu(w)) \\ &= z + F_\nu(w) - F_\mu^{-1}(F_\nu(w)) \\ &= w. \end{aligned}$$

Conversely, suppose $g(z, w) = w$. Then,

$$\begin{aligned} w = g(z,w) &= z + H_\mu(z + H_\nu(w)) = z + F_\mu(z + H_\nu(w)) - (z + H_\nu(w)) \\ &= z + F_\mu(z + F_\nu(w) - w) - (z + F_\nu(w) - w) = w + F_\mu(z + F_\nu(w) - w) - F_\nu(w). \end{aligned}$$

So, $F_\nu(w) = F_\mu(z + F_\nu(w) - w)$, and hence, $F_\mu^{-1}(F_\nu(w)) = z + F_\nu(w) - w$. $\square$

Note that the map $w \to w + F_\mu^{-1}(F_\nu(w)) - F_\nu(w)$ is an non-constant analytic on $\Gamma_{\alpha_2,\beta}$, and $\Gamma_{\alpha_2,\beta}$ is open. By Open Mapping Theorem, the set $\Omega = \{w + F_\mu^{-1}(F_\nu(w)) - F_\nu(w) \mid w \in \Gamma_{\alpha_2,\beta}\} \subseteq \mathbb{C}^+$ is open.

Remark 64. *By Lemma 63, for $z \in \Omega$, g_z has a fixed point in $\mathbb{C}^+$. Now, we are going to show that for every $z \in \mathbb{C}^+$, g_z has a fixed point and that w is an analytic function of z.*

Let $\mathbb{D} = \{z \in \mathbb{C} \mid \|z\| < 1\}$ be the unit disk in the complex plane, and we denote the composition of n copies of f by $f^{\circ n}$. In order to prove the last part of Remark 64, we recall the following two theorems.

Denjoy-Wolff Theorem: Suppose $f : \mathbb{D} \to \mathbb{D}$ is a non-constant analytic function, and f is not an automorphism of $\mathbb{D}$ (i.e., not a Mobius transformation. In other words, not of the form $\lambda(z - \alpha)/(1 - \bar{\alpha}z)$ for some $\alpha \in \mathbb{D}$ and $\lambda \in \mathbb{C}$ with $\|\lambda\| = 1$). If there is a point $w \in \mathbb{D}$ such that $f(w) = w$, then for all $z \in \mathbb{D}$, $f^{\circ n}(z) \to w$. In particular, the fixed point is unique.

Vitali-Porter Theorem: Suppose that $\{f_n\}_{n=1}^\infty$ are sequence of holomorphic functions on $\mathbb{D}$ and $|f_n(z)| \leq M$ for all $n \in \mathbb{N}$, $z \in \mathbb{C}$. If there is an subset $\mathbb{D}' \subset \mathbb{D}$ which has an accumulation point in $\mathbb{D}$ such that $\lim_{n\to\infty} f_n(z)$ exists for all $z \in \mathbb{D}'$, then $\lim_{n\to\infty} f_n(z)$ exists everywhere on $\mathbb{D}$. Furthermore, the convergent is uniform with respect to z in any compact subset of $\mathbb{D}$ and the limit function is analytic on $\mathbb{D}$.

Lemma 65. *Let $g(z,w)$ be as in Definition 61. Then there is a non-constant analytic function $f : \mathbb{C}^+ \to \mathbb{C}^+$ such that for all $z \in \mathbb{C}^+$, $g(z, f(z)) = f(z)$. The analytic function f is unique determined by the fixed point equation.*

Proof. Consider the set

$$\Omega = \{w + F_\mu^{-1}(F_\nu(w)) - F_\nu(w) \mid w \in \Gamma_{\alpha_2,\beta}\} \subseteq \mathbb{C}^+$$

and for $z \in \mathbb{C}^+$, let $g_z : \mathbb{C}^+ \to \mathbb{C}^+$ by $g_z(u) = g(z,u)$. In order to apply Denjoy-Wolff Theorem, consider the map

$$\phi(z) = i\frac{1+z}{1-z} \text{ and } \psi(z) = \frac{z-i}{z+i};$$

ϕ maps $\mathbb{D}$ onto $\mathbb{C}^+$ and ψ maps $\mathbb{C}^+$ onto $\mathbb{D}$. Moreover,

$$\phi(\psi(z)) = i\frac{1+\frac{z-i}{z+i}}{1-\frac{z-i}{z+i}} = i\frac{(z+i)+(z-i)}{(z+i)-(z-i)} = z \text{ and } \psi(\phi(z)) = \frac{i\frac{1+z}{1-z} - i}{i\frac{1+z}{1-z} + i} = \frac{i(1+z)-i(1-z)}{i(1+z)+i(1-z)} = z.$$

Hence, they are inverses of each other.

Let $z \in \Omega$, and $\tilde{g_z} : \mathbb{D} \to \mathbb{D}$ be given by $\tilde{g_z} = \psi \circ g_z \circ \phi$. Since $z \in \Omega$, there is a $w \in \Gamma_{\alpha_2,\beta}$ with $z = w + F_\mu^{-1}(F_\nu(w)) - F_\nu(w)$. Then, $g_z(w) = w$ by Lemma 63. Let $\tilde{w} = \psi(w)$, then $\tilde{w} \in \mathbb{D}$ and we have

$$\tilde{g_z}(\tilde{w}) = \psi(g_z(\phi(\tilde{w}))) = \psi(g_z(w)) = \psi(w) = \tilde{w}.$$

Thus, the map $\tilde{g_z}$ has a fixed point in $\mathbb{D}$. Now, since we have for all $w \in \mathbb{C}^+$,

$$Im(g_z(w)) = Im(z) + Im(H_\mu(z + H_\nu(w))) \geq Im(z).$$

Therefore, g_z is not surjective, and hence g_z is not an automorphism of the upper half plane, and thus, $\tilde{g_z}$ can not be be an automorphism of the disk. Therefore, by Denjoy-Wolff Theorem, $\tilde{g_z}^{\circ n}(\tilde{u}) \to \tilde{w}$ for all $\tilde{u} \in \mathbb{D}$. Thus, $g_z^{\circ n}(u) \to w$ for all $u \in \mathbb{C}^+$.

Now, we define our interates on all of $\mathbb{C}^+$. Let $u_0 = i$ be the initial point. For each $n \in \mathbb{N}$, we define an analytic function $f_n : \mathbb{C}^+ \to \mathbb{C}^+$ by $f_n(z) = g_z^{\circ n}(i)$. We claim that for all $z \in \mathbb{C}^+$, $\lim_{n\to\infty} f_n(z)$ exists. Note that for $z \in \Omega$, we have $\lim_{n\to\infty} f_n(z)$ exists. More precisely, for $z \in \Omega$, we have $f_n(z) = g_z^{\circ n}(i) \to w$.

Now, let $\tilde{\Omega} = \psi(\Omega)$ and $\tilde{f_n} = \psi \circ f_n \circ \phi$. Then, $\tilde{f_n}$ maps $\mathbb{D}$ into $\mathbb{D}$ and for all $\tilde{z} \in \tilde{\Omega}$, $\lim_{n\to\infty} \tilde{f_n}(\tilde{z})$ exists. Hence, by Vitali-Porter Theorem, $\lim_{n\to\infty} \tilde{f_n}(\tilde{z})$ exists for all $\tilde{z} \in \mathbb{D}$. Since $\lim_{n\to\infty} \tilde{f_n}(\tilde{z})$ is not constant on $\tilde{\Omega}$, the Maximum Modulus Principle implies that the limit takes values only in $\mathbb{D}$. Hence, $\lim_{n\to\infty} f_n(z) \in \mathbb{C}^+$ for all $z \in \mathbb{C}^+$. Therefore, we define $f : \mathbb{C}^+ \to \mathbb{C}^+$ by $f(z) = \lim_{n\to\infty} f_n(z)$. By Vitali-Porter Theorem, the convergence is uniform on compact susbsets of $\mathbb{C}^+$ and f is analytic. Since $f_n(z) = g_z^{\circ n}(i)$,

$$g_z(f(z)) = \lim_{n\to\infty} g_z(f_n(z)) = \lim_{n\to\infty} g_z^{\circ(n+1)}(i) = f(z).$$

So, we have $g(z, f(z)) = g_z(f(z)) = f(z)$.

By Denjoy-Wolff Theorem, the function f is uniquely determined by the fixed point equation on the open set Ω; by analytic continuation it is then unique everywhere. $\square$

Theorem 66. *There are analytic functions $\omega_1, \omega_2 : \mathbb{C}^+ \to \mathbb{C}^+$ such that for all $z \in \mathbb{C}^+$,*
(1) $F_\mu(\omega_1(z)) = F_\nu(\omega_2(z))$, and (2) $\omega_1(z) + \omega_2(z) = z + F_\mu(\omega_1(z))$.
The analytic functions ω_1 and ω_2 are uniquely determined by those two equations.

Proof. Let $z \in \mathbb{C}^+$ and $g_z(w) = g(z, w)$. Lemma 65 implies that g_z has a unique fixed point $f(z)$. We define the function $\omega_2(z) = f(z)$ for all $z \in \mathbb{C}^+$, and we define ω_1 by $\omega_1(z) = z + F_\nu(\omega_2(z)) - \omega_2(z)$. Then, ω_1 and ω_2 are analytic on $\mathbb{C}^+$ and $\omega_1(z) + \omega_2(z) = z + F_\nu(\omega_2(z))$.

Let $z \in \Omega$, we have that $z = \omega_2(z) + F_\mu^{-1}(F_\nu(\omega_2(z))) - F_\nu(\omega_2(z))$ by Lemma 63. On the other hand, $z = \omega_2(z) + \omega_1(z) - F_\nu(\omega_2(z))$. Hence, for $z \in \Omega$, we have $\omega_1(z) = F_\mu^{-1}(F_\nu(\omega_2(z)))$. Thus, for all $z \in \Omega$, and hence by analytic continuation for all $z \in \mathbb{C}^+$, we have $F_\mu(\omega_1(z)) = F_\nu(\omega_2(z))$.

Observe the equation (1) and (2), we obtain

$$\begin{aligned} &\omega_1(z) = z + F_\nu(\omega_2(z)) - \omega_2(z) = z + H_\nu(\omega_2(z)) \\ \textit{and} \quad &\omega_2(z) = z + F_\mu(\omega_1(z)) - \omega_1(z) = z + H_\mu(\omega_1(z)). \end{aligned}$$

Thus, $\omega_2(z) = z + H_\mu(z + H_\nu(\omega_2(z))) = g(z, \omega_2(z))$. By Lemma 65, we know that an analytic solution of this fixed point equation is unique. Exchanging the H_μ and H_ν gives in the same way the uniqueness of ω_1. $\square$

We review that if $F : \mathbb{C}^+ \to \mathbb{C}^+$ is analytic, then the theorem of Nevanlinna asserts that there is a unique finite positive Borel measure σ on $\mathbb{R}$ and real numbers a and b, with $b > 0$ such that for $z \in \mathbb{C}^+$

$$F(z) = a + bz + \int_{\mathbb{R}} \frac{1 + tz}{t - z} d\sigma(t).$$

Then given $\alpha > 0$, note that by Lemma 40 we have, for $Im(z) \geq 1$,

$$\left|\frac{1 + tz}{z(t - z)}\right| \leq \frac{1}{|t - z|} + \frac{|t|}{|t - z|} \leq 2\sqrt{1 + \alpha^2}.$$

Then,

$$\frac{F(z)}{z} = \frac{a}{z} + b + \int_{\mathbb{R}} \frac{1 + tz}{z(t - z)} d\sigma(t).$$

For a fixed t we have $(1 + tz)/z(t - z) = (t - 1/z)(t - z) \to 0$ as $z \to \infty$.

Since $|(1 + tz)/z(t - z)|$ is bounded independently of t and z, we apply the Dominated Convergence Theorem to conclude that

$$\frac{F(z)}{z} \to \; b \text{ as } z \to \infty \text{ in } \Gamma_\alpha.$$

Lemma 67. $\lim_{y\to\infty} \omega_1(iy)/iy = \lim_{y\to\infty} \omega_2(iy)/iy = 1.$

Proof. For $i = 1, 2$, we note that $\omega_i : \mathbb{C}^+ \to \mathbb{C}^+$ is analytic, and hence there is a probability measure σ_i such that

$$\omega_i(z) = a_i + b_i z + \int_{\mathbb{R}} \frac{1 + tz}{t - z} d\sigma_i(t).$$

It suffices to show that $b_1 = b_2 = 1$. First, we claim that $\lim_{y\to\infty} \omega_2(iy) = \infty$. That is, for given $\alpha, \beta > 0$ there is $y_0 > 0$ such that $\omega_2(iy) \in \Gamma_{\alpha,\beta}$ whenever $y > y_0$.

Note that $\omega_2(z) = z + H_1(\omega_1(z)) \in z + \mathbb{C}^+$, so we have $Im(\omega_2(z)) > Im(z)$. Then, $b_2 = \lim_{y\to\infty} \omega_2(iy)/(iy) \geq 0$, which implies that $Im(\omega_2(iy))/y \to b_2$ and our inequality $Im(\omega_2(w)) > Im(z)$ implies $b_2 \geq 1$.

Since $Re(\omega_2(iy))/y \to 0$, there is $y_0 > 0$ such that for $y > y_0 \geq \beta$ we have

$$(\frac{Re(\omega_2(iy))}{y})^2 + (\frac{Im(\omega_2(iy))}{y})^2 < (1+\alpha^2)(\frac{Im(\omega_2(iy))}{y})^2.$$

For such a y we have

$$\frac{|\omega_2(iy)|^2}{y^2} < (1+\alpha^2)(\frac{Im(\omega_2(iy))}{y})^2 \Rightarrow |\omega_2(iy)| < \sqrt{1+\alpha^2} Im(\omega_2(iy)).$$

Thus, $\omega_2(iy) \in \Gamma_\alpha$ by Notation 39. Since $Im(\omega_2(iy)) > y > y_0$, we have $\omega_2(iy) \in \Gamma_{\alpha,\beta}$. Thus, $\lim_{y\to\infty} \omega_2(iy) = \infty$.

Since $\omega_1(z) = z + H_\nu(\omega_2(z)) \in z + \mathbb{C}^+$, repeating our argument above, we have $b_1 = \lim_{y\to\infty} \omega_1(iy)/(iy) \geq 1$ and $\lim_{y\to\infty} \omega_1(iy) = \infty$. By Lemma 40, $\lim_{z\to\infty; z\in\Gamma_\alpha} F_\mu(z)/z = 1$, so we have $\lim_{y\to\infty} F_\mu(\omega_1(iy))/\omega_1(iy) = 1$. Therefore, the equation $\omega_1(z) + \omega_2(z) = z + F_\mu(\omega_1(z))$ means that

$$b_1 + b_2 = \lim_{y\to\infty} \frac{\omega_1(iy) + \omega_2(iy)}{iy} = \lim_{y\to\infty} \frac{iy + F_\mu(\omega_1(iy))}{iy} = 1 + \lim_{y\to\infty} \frac{F_\mu(\omega_1(iy))}{\omega_1(iy)} \frac{\omega_1(iy)}{(iy)} = 1 + b_1.$$

Thus, $b_2 = 1$. By the same argument, we have $b_1 = 1$. $\square$

Now, if we let $F = F_\nu \circ \omega_2$, then F maps $\mathbb{C}^+$ to $\mathbb{C}^+$. Moreover,

$$\lim_{y\to\infty} \frac{F(iy)}{iy} = \lim_{y\to\infty} \frac{F_\nu(\omega_2(iy))}{iy} = \lim_{y\to\infty} \frac{F_\nu(\omega_2(iy))}{\omega_2(iy)} \frac{\omega_2(iy)}{iy} = 1.$$

Then applying Theorem 27, we have the following result:

Theorem 68. *Let $F = F_\nu \circ \omega_2$. Then $1/F$ is the Cauchy transform of a probability measure.*

Now, let μ and ν be probability measures on $\mathbb{R}$, then let $F = F_\nu \circ \omega_2 = F_\mu \circ \omega_1$ be as in Theorem 68 and m its corresponding probability measure. By the second part of Theorem 66, we have

$$\omega_1(F^{-1}(1/z)) + \omega_2(F^{-1}(1/z)) - F_\mu(\omega_1(F^{-1}(1/z))) = F^{-1}(1/z).$$

Since $\omega_1(F^{-1}(1/z)) = F_\mu^{-1}(1/z)$ and $\omega_2(F^{-1}(1/z)) = F_\nu^{-1}(1/z)$, we have

$$\begin{aligned}
& F_\mu^{-1}(1/z) + F_\nu^{-1}(1/z) - 1/z = F^{-1}(1/z) \\
\Rightarrow \quad & F_\mu^{-1}(1/z) - 1/z + F_\nu^{-1}(1/z) - 1/z = F^{-1}(1/z) - 1/z \\
\Rightarrow \quad & R_\mu(z) + R_\nu(z) = R_m(z).
\end{aligned}$$

Thus, we achieve our final conclusion:

Theorem 69. *Let μ and ν be two probability measures on $\mathbb{R}$. Then there is a probability measure m on $\mathbb{R}$ with the R-transform R, such that $R_m = R_\mu + R_\nu$.*

Definition 70. Let $\mu \boxplus \nu$ be the probability measure whose Cauchy transform is the reciprocal of F; that is, for which we have $R_{\mu\boxplus\nu} = R_\mu + R_\nu$. We call $\mu \boxplus \nu$ the *free convolution* of μ and ν.

In summary, we consider the probability measures μ and ν with compact support on $\mathbb{R}$ in chatper 3, and we prove that the existence and uniqueness of free convolution (Theorem 34). Then, we try to relax our condition. Consider the case of measures with finite variance in chapter 4, and we give a proof of R-transfrom addition formula under this assumption (Theorem 51.) Finally, we discuss the most general case in chapter 5; that is, we consider probability measures on $\mathbb{R}$ that may not have any moments. It is amazing that the R-transform addition formula still hold (Theorem 69), and it is the final conclusion in the thesis.

Bibliography

[1] U. Haagerup. *On Voiculescu's R- and S-transform for free non-commuting random variables*, in: D. Voiculescu (Ed.), Free Probability Theory, Fields Institute Communications, vol. 12, 1997, pp. 127-148.

[2] A. Nica and R. Speicher, *Lectures on the combinatorics of free probability*, Cambridge University Press, 2006.

[3] D. Voiculescu. *Symmetries of some reduced free product C*-algebras*, in: Operator Algebras and Their Connection with Topology and Ergodic Theory, Lecture Notes in Math., vol. 1132, Springer-Verlag (1985), pp. 556-588.

[4] J. A. Mingo and R. Speicher. *Free probability and random matrices*, Springer-Verlag (unpublished).

[5] Anderson, Greg W., Alice Guionnet, and Ofer Zeitouni. *An introduction to random matrices*, Cambridge University Press, 2010.

[6] D. Vociulescu. *Limit laws for random matrices and free products*. Invent. Math. 104 (1991), 201-220.

[7] H. Maassen. *Addition of freely independent random variables*. J. Funct. Anal. 106 (1992), no. 2, 409-438.

[8] H. Bercovici and D. Voiculescu. *Free convolution of measures with unbounded support*. Indiana Univ. Math. J. 42 (1993), no. 3, 733-773.

[9] S. T. Belinschi and H. Bercovici. *A new approach to subordination results in free probability*. J. Anal. Math. 101 (2007), 357-365.